Integrating physiological and psychoacoustic phenomena into modeling of binaural hearing

von

Jonas Klug

geboren am 17. Oktober 1991
in Hildesheim (Deutschland)

Erstgutachter: Prof. Dr. Mathias Dietz
Zweitgutachter: Prof. Dr. Volker Hohmann
Dritte Prüferin: Prof. Dr. Christine Köppl

Tag der Disputation: 10. Oktober 2023

Acknowledgments

This thesis was written during my time as a PhD student at the Physiology and Modeling of Auditory Perception Division of the Department of Medical Physics and Acoustics at the Carl von Ossietzky University of Oldenburg. I'd like to thank the many people who contributed in various ways to make this work possible.

Most importantly, I'd like to thank my supervisor and group leader Prof. Mathias Dietz for creating an open working environment and for his continuous support, scientific advice and discussions. He not only introduced me to jet skiing, but also got me excited about hearing research and connected me with many researchers around the world. I am thankful for his analytical and critical approach to science, which I gladly adopt, and which has driven us in our search for the truth.

I am grateful that I got to know Prof. Laurel Carney. Laurel's advice and broad perspective on auditory research greatly enriched my experience as a PhD student. I am also grateful to Prof. Volker Hohmann, whose feedback has been an asset to my research. Of particular note are his comments about the level dependence of auditory processing explored in the first paper included in this thesis. Furthermore I thank Prof. Christine Köppl and Dr. Amarins Heeringa for our work on the auditory nerve. Dr. Sean Anderson, it was awesome to have you as a mate: whether it was discussions about the CN, ARO session or road-trip - epic time.

Special thanks to all current and former members of the Phymo group for the very nice and relaxed working atmosphere, the many scientific and non-scientific discussions and of course great events like Kohltours or the Florida-Trip. Special thanks to my office mates Swantje Hansen and Dr. Henri Pöntynen, you helped a lot in finishing this project. We also had a lot of fun turning our office into a jungle and tasting specialities from different regions. Thank you Henri for being a friend! I much enjoyed the time with Lisa Schmors and Helen Heinermann who introduced me to the world of neuroscience. Lisa, we did a great job on our paper. Helen, I am grateful for our time together, especially our road trip in California was a highlight. Thanks to Dr. Jörg Encke and Dr. Go Ashida, both of you were always ready to explain complex things. Thank you for sharing your great wealth of knowledge about neurons and their connections. Thank you Jörg for the Espresso-trick. The PhD student group was also important during my time as a PhD student, as it was always a place to share worries and frustrations, but also to celebrate successes. Thank you Sven Herrmann for many coding discussions and cooperation. Anna Dietze, it is always a pleasure to have you around and I am grateful for our time in the sauna, garden and in the mountains. Bernhard Eurich, thank you for the hours that we were able to spend together, some of which were refined with whisky. Rebecca Felsheim, thank you for putting coding on the agenda again. Thank you Dr. Hongmei Hu for introducing me to CI

processing. I appreciate that we both enjoy the little things.

Zusammenfassung

Für das räumliche Hören sind die sogenannten *Binauralen Cues* wichtig. Durch interaurale Zeit- und Pegel-Differenzen (ITD und ILD) zwischen den beiden Ohren ist es möglich Geräuschquellen zu orten. Für die Sprachverständlichkeit während eines Störgeräusches, die oft eine Herausforderung für Hörgeschädigte darstellt, ist das binaurale Hören von enormer Bedeutung.

Diese Doktorarbeit befasst sich mit den grundlegenden Aspekten des binauralen Hörens und seiner Modellierung. Dabei wird großer Wert auf die physiologische Plausibilität der Modellierung gelegt, da nur so Rückschlüsse auf mögliche Ursachen von Hörschwierigkeiten und in der Folge auf Konzepte für Hörgeräte-Algorithmen gezogen werden können. Daher konzentriert sich diese Forschung auf die Integration psychoakustischer und physiologischer Erkenntnisse in die Modellierung des binauralen Hörens.

Ein Standardmodell für binaurales Hören basiert auf einer internen Kreuzkorrelation zwischen den Signalen des linken und rechten Ohres. Dieses Prinzip erfordert zwei Reihen von entgegengesetzten Verzögerungsgliedern, eine sogenannte Delay-Line. Es hatte sich gezeigt, dass dieses Delay-Line-Modell nicht mit der Physiologie von Säugetieren übereinstimmt und einige psychoakustische Datensätze nicht erklären kann. Als Alternative wurde das Opponent-Channel-Modell mit nur zwei gegenläufigen Kanälen diskutiert, das als physiologisch plausibler angesehen wird, aber bisher weniger Daten erklären konnte.

In Kapitel 3 wird ein Opponent-Channel-Modell entwickelt, das viele Daten zur ILD und Einhüllenden-ITD Wahrnehmung erklären kann. Dabei wird eine umfassendere Erklärung geboten, indem psychoakustische Phänomene mit physiologischen Erkenntnissen in Einklang gebracht werden. Wichtige Erkenntnisse sind hier, dass die Modell-Neuronen ITDs und ILDs gleichzeitig kodieren und dass bei verschiedenen Pegeln verschiedene Frequenzkanäle die binauralen Informationen enthalten.

Ein weiterer Schwerpunkt liegt auf der Wahrnehmung der Feinstruktur-ITD und ihrer Reduktion mit zunehmender Frequenz. Die Erklärung dieses psychoakustischen Datensatzes stellt bisher eine Herausforderung für beide Modellierungsansätze dar. Am Hörnerv nimmt die Repräsentation der Phase (Phase-Locking) zu höheren Frequenzen hin ab. Die gängige Theorie, dass diese Abnahme für die Reduktion der ITD-Wahrnehmung verantwortlich ist, wurde abgeschwächt: Die Analyse mehrerer Hörnerv-Datensätze in Kapitel 4 zeigt eine Reduktion des Phase-Lockings um $\leq 60\,\mathrm{dB/Dekade}$, was flacher ist als bisher angenommen. Im Gegensatz dazu wird in Kapitel 5 in einer psychoakustischen Studie die Abnahme der ITD-Empfindlichkeit genauer gemessen. Die Analyse ergibt eine Reduktion um 153-260 dB/Dekade. Es wurde auch gezeigt, dass die höchste Empfindlichkeit bei einer interauralen Phasendifferenz von etwa $0,8\pi$ im Bogenmaß liegt und nicht bei π, wie es eine einfache Modellierung erwarten ließe. In Kapitel 6 wird eine Reihe von Hypothesen zur Erklärung der steilen Reduktion der ITD-Wahrnehmung

vorgestellt und diskutiert. Dabei wird zum einen ein Experiment vorgestellt, das die Hypothese der dominanten Region schwächt und zum anderen gezeigt, dass der synaptische Filter zumindest theoretisch eine hohe Reduktion erzeugen kann.

Abstract

Binaural cues are important for spatial hearing. Through interaural time and level differences (ITD and ILD) between the two ears it is possible to locate sound sources. Binaural hearing is of great importance for speech intelligibility in noise, which is often a challenge for hearing impaired people.

This thesis deals with the fundamental aspects of binaural hearing and its modeling. Great emphasis is placed on the physiological plausibility of the modeling, as this is the only way to draw conclusions about possible causes of hearing loss and concepts for hearing aid algorithms. Therefore, this research focuses on the integration of psychoacoustic and physiological knowledge into the modeling of binaural hearing.

A standard model of binaural detection is based on internal cross-correlation between the signals from the left and right ear. This principle requires two rows of opposing delay elements, called a delay line. This delay-line model was found to be incompatible with mammalian physiology and unable to explain some psychoacoustic data. As an alternative, the opponent-channel model with only two opposing channels was discussed, which is considered physiologically more plausible, but could explain less data.

In chapter 3, an opponent-channel model is developed that can explain much of the data on ILD and envelope-ITD perception. It provides a more comprehensive explanation by reconciling psychoacoustic phenomena with physiological findings. Important findings are that the model neurons encode ITDs and ILDs simultaneously, and that different frequency channels contain the binaural information at different levels.

Another focus is the perception of fine structure ITD and its strong decrease towards higher frequencies. Explaining this psychoacoustic data has been a challenge for both modeling approaches.

The common theory that the decrease in synchrony towards higher frequencies at the auditory nerve is responsible for this decrease has been weakened: The analysis of several auditory nerve data sets in chapter 4 shows a decrease in synchrony of $\leq 60\,\mathrm{dB/decade}$, which is shallower than previously assumed. In contrast, in a psychoacoustic study in chapter 5 the decline in ITD sensitivity is measured more precisely, and the analysis shows a decline of 153-260 $\mathrm{dB/decade}$. It was also shown that the highest sensitivity occurs at an interaural phase difference of about 0.8π in radians, and not at π as simple modeling would suggest. In chapter 6 a number of hypotheses to explain the steep gradient of ITD perception are presented and discussed. On the one hand, an experiment is presented that weakens the dominant region hypothesis, and on the other hand, it is shown that the synaptic filter can, at least theoretically, produce a high gradient.

Glossary

Acronyms and Abbreviations

dB   decibel

dec   decade

oct   octave

rad.   radians

sps   spikes per second

AFC   Alternative Forced-Choice

AM   Amplitude Modulation

AN   Auditory Nerve

AVCN   Anteroventral CN

BM   Basilar Membrane

BW   Bandwidth

CF   Characteristic Frequency

CI   Cochlear Implant

CN   Cochlear Nucleus

EE   Excitatory-Excitatory

EEG   Electroencephalography

EI   Excitatory-Inhibitory

EPSP   Excitatory Postsynaptic Potential

ENV   Envelope

ERB   Equivalent Rectangular Bandwidth

GBC   Globular Bushy Cell

GN   Gaussian Noise

HL   Hearing Level

IC	Inferior Colliculus
IHC	Inner Hair Cell
ILD	Interaural Level Difference
IPD	Interaural Phase Difference
IPSP	Inhibitory Postsynaptic Potential
ITD	Interaural Time Difference
ITD_{TFS}	Temporal Fine Structure ITD
ITD_{ENV}	Envelope ITD
LSO	Lateral Superior Olive
MNTB	Medial Nucleus of the Trapezoid Body
MSO	Medial Superior Olive
OHC	Outer Hair Cell
PSTH	Peristimulus Time Histogram
RC	Resistor-Capacitor
RMSE	Root-Mean-Square Error
SAM	Sinusoidally Amplitude Modulated
SBC	Spherical Bushy Cell
SI	Synchrony Index
SOC	Superior Olivary Complex
SPL	Sound Pressure Level
TFS	Temporal Fine Structure
VAF	Variance Accounted For
VS	Vector Strength

Fixed symbols

a amplitude of SAM

b y-axis intercept

$b(k, N, p)$ binomial probability mass function

$c(\tau)$ cross-correlation

d' d prime

$\hat{d}'$ prediction of d'

f frequency

f_b knee point

f_c carrier frequency

f_0 corner frequency of a first order filter

f_corner corner frequency

f_m modulation frequency

g gain

g_m magnitude of IPD dependence

i imaginary unit / condition index

j frequency channel index / data point index

k spike index / number of correct choices

k_c convolution kernel (from synaptic filter)

l likelihood

m decline (negative slope)

m_d modulation depth

n exponent for raised sine / filter order

p probability

$p(\tau)$ density function (centroid) across best ITD

p_c proportion correct

r correlation coefficient

r_s spontaneous rate

t time

t_abs absolute refractory time

t_rel relative refractory time

v VS

$\overline{y}$ mean of the observed pointer ILDs

y_i observed pointer ILD for the ith condition

$\hat{y}_i$ predicted pointer ILD for the ith condition

$G(f)$ filter gain across frequency

H magnitude of α-function

J number of data points

K number of spikes

M_{ex} number of excitatory inputs

M_{inh} number of inhibitory inputs

N number of frequency channels / number of times measured

NH Normal Hearing

$R_{j\text{L}}$ rate of an EI-model-neuron in the left hemisphere at the respective $\text{CF}(j)$

$R_{j\text{R}}$ rate of an EI-model-neuron in the right hemisphere at the respective $\text{CF}(j)$

T refractory period

W_{ex} excitatory window duration

W_{inh} inhibitory window duration

α_k angle of the k-th spike, as a unit vector corresponding to spike time within cycle

δ inhibitory gain

ϵ RMSE

θ response threshold

μ mean of the distribution of the internal representation of a stimulus

ρ subject specific scaling factor

σ standard deviation of the distribution of the internal representation of a stimulus

σ_0 σ for IPD $= 0$

$\overline{\sigma}(f)$ mean value of σ over ΔIPD across frequency

τ best ITD / time constant of α-function

τ_{e} time constant of EPSP

τ_{i} time constant of IPSP

$\overline{\tau}(\Delta t)$ centroid measure

ψ VAF

Δt ITD

Δt_{i} timing offset between EPSP and IPSP

ΔR_j hemispheric rate difference at a respective $\text{CF}(j)$

$\Delta \overline{R}$ mean spike rate difference between the left and right hemisphere

Contents

Contents

1. Introduction

Binaural hearing allows localization of sound sources in space without prior knowledge about their location or visual cues. This is essential for identifying the directions of important sounds such as someone calling your name or potential dangers such as traffic. For animals, sounds of interest may come from prey, mates or predators. In addition to localization, the brain can organize auditory scenes with multiple sound sources into perceptual representations of the different sources, allowing individuals to focus on a single target source in competing background noise. This ability is critical for communication.

The dominant cues for estimating the direction in the horizontal plane are the differences in arrival time and intensity between the left and right ear (Strutt, 1907). They are called binaural cues. Since the head forms an acoustic obstacle to the traveling waves (the so-called head shadow effect), an interaural level difference (ILD) emerges (Blauert, 1996). At low frequencies, where the wavelength of sound is larger than the dimensions of the head, this effect is marginal, while at high frequencies ILDs up to 20 - 30 dB are possible for humans (Kayser *et al.*, 2009; Mayo and Goupell, 2020).

Additionally, if a sound is presented on one side, the sound wave arrives at different times at the two ears. This leads to an interaural time difference (ITD). For a sound wave arriving at both ears, the ITD depends on the source position and the size of the head. ITD is maximum for a source position fully to the right (+ 90°), or fully to the left (- 90°) of the head. For this angular range (± 90°), the range of naturally occurring ITDs for human heads spans about ±700 μs (Blauert, 1996). A distinction is made between a delay of the temporal fine structure (ITD$_{\text{TFS}}$) and a delay of the envelope (ITD$_{\text{ENV}}$). From the ITD, the interaural phase difference (IPD) can be calculated by multiplying the ITD with the respective frequency f:

$$\text{IPD} = 2\pi \ f \ \text{ITD}. \tag{1.1}$$

With the multiplication of 2π, the IPD is given in radians.

1. Introduction

The auditory system is studied using three mutually informative methodologies (1) psychoacoustic, (2) physiology, and (3) modeling (Figure 1.1). Usually, the hypothesis for a psychoacoustic and/or physiological experiment arises from a modeling idea, or models can be refined based on behavioral data. A good model can be the crucial factor for getting a deeper understanding of the neural mechanisms underlying behavioral or physiological outcomes. This methodological triad is important for developing a good understanding of the coding of binaural hearing and is illustrated below.
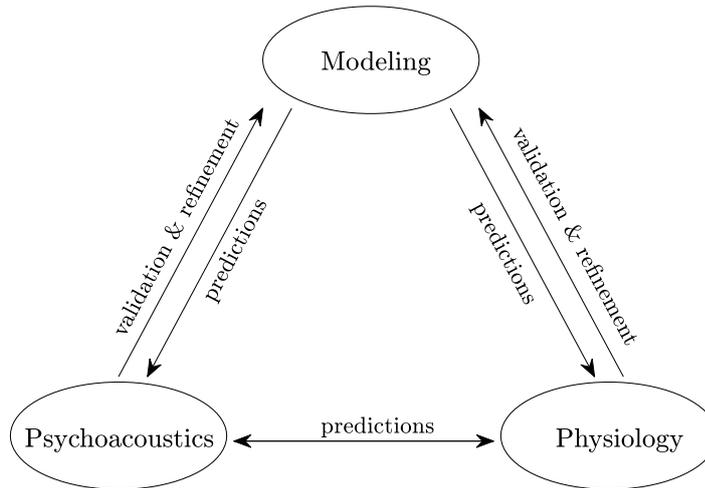
Figure 1.1.: Triad of methods of this research.

Psychoacoustic studies have shown that the ITD cue plays a dominant role for source localization at low frequencies and the ILD cue at high frequencies (Strutt, 1907; Macpherson and Middlebrooks, 2002). Although this *duplex theory* of binaural hearing has been a widely accepted conceptualization of human sound localization for more than a century, the neural mechanisms underlying binaural processing remain poorly understood. A more comprehensive model requires a better understanding of the neural processes underlying the encoding of binaural information in the auditory system. It has been suggested that ITD and ILD are processed in separate pathways in the auditory system (cf., Tollin, 2003) due to the differences in their initial encoding requirements and in accordance with the duplex theory. Therefore, it has long been thought that ILDs and ITDs are likely to be encoded in anatomically separate but parallel pathways in the auditory system (Grothe *et al.*, 2010; Owrutsky *et al.*, 2021).

Anatomy and physiology provide important insights into the functioning of binaural interaction. Therefore, an overview of auditory physiology is provided in Chapter 2. Details about the pre-processing performed by the auditory periphery can be found in Section 2.1. The first anatomical structure where signals from the right and the left side are combined is the superior olivary complex (SOC), located in both hemispheres of the auditory brainstem. This is described in Section 2.2. Two principal nuclei are part of the SOC: the medial superior olive (MSO) and the lateral superior olive (LSO). From the details on the SOC in Section 2.2 one could conclude, that the MSO is sensitive to ITDs in the temporal fine structure (ITD_{TFS}), whereas the LSO is mainly sensitive to ILDs and ITDs in the envelope (ITD_{ENV}) (Grothe *et al.*, 2010).

The first model of ITD encoding was developed by Jeffress (1948), who hypothesized the exis-

tence of coincidence detecting neurons and proposed an array of such detectors, arranged along a neural delay line. Each neuron in this delay line would respond maximally if the neuron-specific delay of the inputs compensates the external ITD. The ITD for which the highest firing rate of a respective neuron is achieved is called best ITD. In such a network, the position of the coincidence detector with the highest activity is a measure for the ITD. The operation of this mechanism is closely related to cross-correlation of the left and right ear signals, where the delay between the two input signals is determined by the position of the highest peak. Based on this cross-correlation (delay-line) model and its extensions (e.g. Colburn, 1977; Stern and Colburn, 1978; Stern and Shear, 1996), many psychoacoustic data (e.g. Sayers, 1964; Culling, 2007; Bernstein and Trahiotis, 2012) have been modeled and successfully explained. With the physiological evidence for delay-lines in birds (Carr and Konishi, 1988, 1990), the model seems consistent at least with avian physiology.

Nevertheless, more recent studies on mammalian inferior colliculus (IC) and MSO neurons in guinea pigs (McAlpine *et al.*, 2001) and gerbils (Brand *et al.*, 2002) have shown that these neurons are tuned to a wide range of ITDs, with most of them having their best ITDs either at the extreme ends or beyond the physiological range of the animal. In Marquardt and Mcalpine (2007), most of the IC neurons have their best ITD around $\pi/3$ radians. This observation is inconsistent with a delay line mechanism, which would require a large number of neurons with their best ITDs distributed across the physiological range.

An alternative mechanism for encoding ITD involves comparing the firing rates of nuclei in the two hemispheres. This approach has been called the count-comparison (Colburn and Durlach, 1978), the hemifield (Stecker *et al.*, 2005) or the opponent-channel (Magezi and Krumbholz, 2010) model. While some human psychoacoustic data has been successfully predicted using such opponent-channel models (Dietz *et al.*, 2011; Takanen *et al.*, 2014; Encke and Hemmert, 2018), the internal representation of spatial perception and the process by which it is analyzed is still debated (McAlpine *et al.*, 2001; Bernstein and Trahiotis, 2017; Encke and Dietz, 2022). Due to its simplicity and ease of implementation, the delay-line model is efficient when executed as a cross-correlation and has gained popularity in technical applications. However, to gain a deeper understanding of the human auditory system and its possible pathologies, the physiological plausibility of the internal representation is important.

1. Introduction

The aim of this thesis is to expand the current understanding of human binaural hearing. An important goal of the work is to develop an opponent-channel model that accounts for the behavioral data set of Bernstein and Trahiotis (2012). This eminent data set for human binaural hearing contains 960 data points with respect to ILD and ITD_{ENV}. The fact that this extensive data set can be explained very well by a delay-line model has been considered a strong indication that the human auditory binaural system is based on such a mechanism. However, the opponent channel model is considered more physiologically plausible, but has yet to be validated as an explanation for many data (Encke and Dietz, 2022).

Chapter 3 (based on Klug *et al.*, 2020) shows that an opponent-channel model not only explains this data set as well as similar data sets (Bernstein and Trahiotis, 2003; Dietz *et al.*, 2015), but also predicts physiologic measurements of Joris and Yin (1995). This is an important result, as the opponent-channel model is better in-line with current knowledge about mammalian auditory physiology than the delay-line based approaches. For the binaural interaction stage a spiking LSO model developed by Ashida *et al.* (2016) was used. This modeling study highlighted the importance of the across frequency processing and the pre-processing in the auditory periphery (see Section 2.1).

Decreasing sensitivity to ITD_{TFS} (e.g. Brughera *et al.*, 2013) and ITD_{ENV} (e.g. Bernstein and Trahiotis, 2002) to higher frequencies or modulation frequencies is an important characteristic of behavioral and physiological data (cf. Figure 2.7 B and Figure 2.8 D). These properties are critical and especially the steep decline in ITD_{TFS} sensitivity is examined here in several chapters. The classic explanation for this decline is the decline of AN phase locking across frequency (e.g. Verschooten *et al.*, 2019).

Chapter 4 (based on Klug *et al.*, 2023) presents an updated set of low-pass filter parameters that accounts for several AN phase locking data sets, revealing that phase locking declines with frequency at much lower rate ($\leq 18\,\text{dB/oct}$ or $\leq 60\,\text{dB/dec}$) than previously assumed. On the other hand, **Chapter 5** (based on Klug and Dietz, 2022) presents an experiment to reveal the exact decline of IPD_{TFS} sensitivity across frequency. This information was hidden by the limited threshold measurements by Brughera *et al.* (2013). Nevertheless, the decline in IPD sensitivity of 46-78 dB/oct (153-260 dB/dec) is in line with this previous data set. This dramatic difference in slope between the AN phase locking and the ITD_{TFS} sensitivity reveals a gap in our understanding of the binaural processing of temporal differences.

Possible mechanisms accounting for the steep decline are discussed in **Chapter 6**. The interaction of excitatory and inhibitory post-synaptic potentials (EPSPs and IPSPs) proved to be a key principle in accounting for human psychophysical data (Chapter 3). The conceptualization of an EPSP as synaptic low-pass filter offers an explanation to the decline of ITD_{ENV} sensitivity towards higher modulation frequencies. The synaptic filter may be the key factor causing the large slope in ITD_{TFS} sensitivity shown in Klug and Dietz (2022).

1.1. References

Ashida, G., Kretzberg, J., and Tollin, D. J. (**2016**). "Roles for coincidence detection in coding amplitude-modulated sounds," PLOS Computational Biology **12**(6), 1–27, doi: `10.1371/journal.pcbi.1004997`.

Bernstein, L. R., and Trahiotis, C. (**2002**). "Enhancing sensitivity to interaural delays at high frequencies by using "transposed stimuli"," J. Acoust. Soc. Am. **112**(3), 1026–1036, doi: `10.1121/1.1497620`.

Bernstein, L. R., and Trahiotis, C. (**2003**). "Enhancing interaural-delay-based extents of laterality at high frequencies by using 'transposed stimuli'," The Journal of the Acoustical Society of America **113**(6), 3335–3347, doi: `10.1121/1.1570431`.

Bernstein, L. R., and Trahiotis, C. (**2012**). "Lateralization produced by interaural temporal and intensitive disparities of high-frequency, raised-sine stimuli: Data and modeling," The Journal of the Acoustical Society of America **131**(1), 409–415, doi: `10.1121/1.3662056`.

Bernstein, L. R., and Trahiotis, C. (**2017**). "An interaural-correlation-based approach that accounts for a wide variety of binaural detection data," The Journal of the Acoustical Society of America **141**(2), 1150–1160, doi: `10.1121/1.4976098`.

Blauert, J. (**1996**). *Spatial Hearing: The Psychophysics of Human Sound Localization* (The MIT Press).

Brand, A., Behrend, O., Marquardt, T., McAlpine, D., and Grothe, B. (**2002**). "Precise inhibition is essential for microsecond interaural time difference coding," Nature **417**(6888), 543–547, doi: `10.1038/417543a`.

Brughera, A., Dunai, L., and Hartmann, W. M. (**2013**). "Human interaural time difference thresholds for sine tones: The high-frequency limit," J. Acoust. Soc. Am. **133**(5), 2839–2855, doi: `10.1121/1.4795778`.

Carr, C., and Konishi, M. (**1990**). "A circuit for detection of interaural time differences in the brain stem of the barn owl," Journal of Neuroscience **10**(10), 3227–3246, doi: `10.1523/JNEUROSCI.10-10-03227.1990`.

Carr, C. E., and Konishi, M. (**1988**). "Axonal delay lines for time measurement in the owl's brainstem," Proceedings of the National Academy of Sciences **85**(21), 8311–8315, doi: `10.1073/pnas.85.21.8311`.

Colburn, H. S. (**1977**). "Theory of binaural interaction based on auditory-nerve data. ii. detection of tones in noise," The Journal of the Acoustical Society of America **61**(2), 525–533, doi: `10.1121/1.381294`.

Colburn, H. S., and Durlach, N. I. (**1978**). "Models of binaural interaction," in *Handbook of perception: Hearing*, edited by E. C. Carterette and M. P. Friedman, **4** (Academic Press, New York), pp. 467–518.

1. Introduction

Culling, J. F. (**2007**). "Evidence specifically favoring the equalization-cancellation theory of binaural unmasking," The Journal of the Acoustical Society of America **122**(5), 2803–2813, doi: `10.1121/1.2785035`.

Dietz, M., Ewert, S. D., and Hohmann, V. (**2011**). "Auditory model based direction estimation of concurrent speakers from binaural signals," Speech Communication **53**(5), 592–605, doi: `10.1016/j.specom.2010.05.006` perceptual and Statistical Audition.

Dietz, M., Klein-Hennig, M., and Hohmann, V. (**2015**). "The influence of pause, attack, and decay duration of the ongoing envelope on sound lateralization," The Journal of the Acoustical Society of America **137**(2), EL137–EL143, doi: `10.1121/1.4905891`.

Encke, J., and Dietz, M. (**2022**). "A hemispheric two-channel code accounts for binaural unmasking in humans," Communications Biology **5**(1), 1122, doi: `10.1038/s42003-022-04098-x`.

Encke, J., and Hemmert, W. (**2018**). "Extraction of inter-aural time differences using a spiking neuron network model of the medial superior olive," Frontiers in Neuroscience **12**, 140, doi: `10.3389/fnins.2018.00140`.

Grothe, B., Pecka, M., and McAlpine, D. (**2010**). "Mechanisms of sound localization in mammals," Physiological Reviews **90**(3), 983–1012, doi: `10.1152/physrev.00026.2009`.

Jeffress, L. (**1948**). "A place theory of sound localization," Journal of comparative and physiological psychology **41**(1), 35—39, doi: `10.1037/h0061495`.

Joris, P. X., and Yin, T. C. (**1995**). "Envelope coding in the lateral superior olive. I. Sensitivity to interaural time differences," Journal of Neurophysiology **73**(3), 1043–1062, doi: `10.1152/jn.1995.73.3.1043`.

Kayser, H., Ewert, S., Anemüller, J., Rohdenburg, T., Hohmann, V., and Kollmeier, B. (**2009**). "Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses," Eurasip Journal on Advances in Signal Processing **2009**, 10, doi: `10.1155/2009/298605`.

Klug, J., and Dietz, M. (**2022**). "Frequency dependence of sensitivity to interaural phase differences in pure tones," The Journal of the Acoustical Society of America **152**(6), 3130–3141, doi: `10.1121/10.0015246`.

Klug, J., Encke, J., and Dietz, M. (**2023**). "Characterization of the decline in the auditory nerve phase locking at high frequencies," JASA Express Letters **3**(7), 074403, doi: `10.1121/10.0020267`.

Klug, J., Schmors, L., Ashida, G., and Dietz, M. (**2020**). "Neural rate difference model can account for lateralization of high-frequency stimuli," The Journal of the Acoustical Society of America **148**(2), 678–691, doi: `10.1121/10.0001602`.

Macpherson, E. A., and Middlebrooks, J. C. (**2002**). "Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited," The Journal of the Acoustical Society of America **111**(5), 2219–2236, doi: `10.1121/1.1471898`.

Magezi, D. A., and Krumbholz, K. (**2010**). "Evidence for opponent-channel coding of interaural time differences in human auditory cortex," Journal of Neurophysiology **104**(4), 1997–2007, doi: `10.1152/jn.00424.2009` pMID: 20702739.

Marquardt, T., and Mcalpine, D. (**2007**). "A π-limit for coding ITDs: Implications for binaural models," in *Hearing – From Sensory Processing to Perception*, edited by B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, and J. Verhey, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 407–416, doi: `10.1007/978-3-540-73009-5_44`.

Mayo, P. G., and Goupell, M. J. (**2020**). "Acoustic factors affecting interaural level differences for cochlear-implant users," The Journal of the Acoustical Society of America **147**(4), EL357–EL362, doi: `10.1121/10.0001088`.

McAlpine, D., Jiang, D., and Palmer, A. R. (**2001**). "A neural code for low-frequency sound localization in mammals," Nature neuroscience **4**(4), 396–401, doi: `10.1038/86049`.

Owrutsky, Z. L., Benichoux, V., and Tollin, D. J. (**2021**). "Binaural hearing by the mammalian auditory brainstem: Joint coding of interaural level and time differences by the lateral superior olive," in *Binaural Hearing*, edited by R. Y. Litovsky, M. J. Goupell, R. R. Fay, and A. N. Popper (Springer International Publishing, Cham), pp. 113–144, doi: `10.1007/978-3-030-57100-9_5`.

Sayers, B. M. (**1964**). "Acoustic-image lateralization judgments with binaural tones," J. Acoust. Soc. Am. **36**(5), 923–926, doi: `10.1121/1.1919121`.

Stecker, G. C., Harrington, I. A., and Middlebrooks, J. C. (**2005**). "Location coding by opponent neural populations in the auditory cortex," PLOS Biology **3**(3), doi: `10.1371/journal.pbio.0030078`.

Stern, R. M., and Colburn, H. S. (**1978**). "Theory of binaural interaction based on auditory-nerve data. IV. A model for subjective lateral position," J. Acoust. Soc. Am. **64**(1), 127–140, doi: `10.1121/1.381978`.

Stern, R. M., and Shear, G. D. (**1996**). "Lateralization and detection of low-frequency binaural stimuli: Effects of distribution of internal delay," J. Acoust. Soc. Am. **100**(4), 2278–2288, doi: `10.1121/1.417937`.

Strutt, J. W. (**1907**). "XII. On our perception of sound direction," The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **13**(74), 214–232, doi: `10.1080/14786440709463595`.

Takanen, M., Santala, O., and Pulkki, V. (**2014**). "Visualization of functional count-comparison-based binaural auditory model output," Hearing Research **309**, 147–163, doi: `10.1016/j.heares.2013.10.004`.

Tollin, D. J. (**2003**). "The lateral superior olive: A functional role in sound source localization," The Neuroscientist **9**(2), 127–143, doi: `10.1177/1073858403252228` pMID: 12708617.

1. Introduction

Verschooten, E., Shamma, S., Oxenham, A., Moore, B., Joris, P., Heinz, M., and Plack, C. (**2019**). "The upper frequency limit for the use of phase locking to code temporal fine structure in humans: A compilation of viewpoints," Hearing Research **377**, doi: `10.1016/j.heares.2019.03.011`.

2. Processing in the auditory pathway

In this chapter, a brief summary is given of the key stages of the mammalian auditory pathway that play a role in sound localization. Similar to the previous section, the focus will be on interaural level difference (ILD) and interaural phase difference (IPD), the spatial cues that are relevant to the scope of this thesis. The focus here is on the processing of pure tones and sinusoidally amplitude modulated (SAM) tones. For a more comprehensive overview, readers may refer to the available literature (e.g. Grothe *et al.*, 2010; Plack, 2013).

2.1. Peripheral processing: From sound to auditory nerve spikes

The peripheral auditory system is responsible for gathering acoustic signals and transforming them into nervous system activity. The peripheral system consists of three main parts: the outer ear, middle ear, and inner ear. The outer ear includes the pinna, which collects and amplifies sound waves before directing them to the middle ear

The middle ear consists of the ossicular chain, which transmits the sound wave from the tympanic membrane to the inner ear. The inner ear consists of the semicircular canals, the vestibule and the cochlea. The cochlea is a spiral-shaped structure responsible for converting mechanical sound waves into neural signals that are transmitted to the brain. The basilar membrane (BM) in the cochlea responds to the mechanical motion of the fluids in the cochlea, causing a wave to travel along the membrane from the base to the apex (von Békésy, 1970). The BM is wider and more flexible at the apical end and narrower and stiffer at the basal end. This change in mechanical properties results in a tonotopic mapping along the membrane, so that low frequency sounds cause a maximum membrane displacement at the apical end and high frequencies at the basal end of the BM. Because different sections of the BM respond to different sound frequencies in a selective manner, the BM is often conceptualized as a bank of band-pass filters. For modeling the peripheral filtering in the cochlea in computational auditory models the implementation of the Gammatone filter from Hohmann (2002) is widely used (e.g. Jürgens and Brand, 2009; Klein-Hennig *et al.*, 2011; Søndergaard and Majdak, 2013; Eurich *et al.*, 2022). The bandwidth of this auditory filters is often expressed as the equivalent rectangular bandwidth (ERB) (Glasberg and Moore, 1990).

The outer hair cells (OHC) provide some amplification and thus contribute significantly to the frequency selectivity of the BM (Ashmore, 1987). OHCs are able to contract and expand in response to different sound levels. This so called "cochlear amplifier" increases the amplitude of BM displacement for low levels and enhances the sensitivity of the inner ear. As the sound level increases, the contraction of the OHCs effectively reduces their mechanical sensitivity. This compression mechanism allows the cochlea to handle a wide range of sound intensities without saturating the auditory system or damaging the delicate structures within the inner ear (Robles and Ruggero, 2001).

2. Processing in the auditory pathway

The inner hair cells (IHCs), located between the basilar and tectorial membrane, are stimulated by the bending of the stereocilia on top of the cell bodies (Hudspeth, 1985). This bending is caused by shear forces between the two membranes. The movement of the stereocilia causes mechanically gated ion channels to open. A receptor potential is induced in the hair cells by positively charged potassium ions flowing into them. Because the ion channels open only when the stereocilia are bent towards one direction, the outputs of the IHCs are half-wave rectified versions of the input signals (Breebaart *et al.*, 2001a). The speed of the stereocilia movement is limited and therefore introduces low-pass filtering into the system (Meddis *et al.*, 2010). The receptor potential induction process involves the activation of voltage-gated calcium channels and the subsequent release of neurotransmitters into the gap between the IHCs and the auditory afferent neurons. This leads to stimulation of the auditory nerve (AN) and the generation of an action potential. The stimulation of AN fibers by IHCs is tonotopically organized, meaning that each nerve fiber is characterized by the frequency it is most sensitive to (Greenwood, 1961). This frequency is called characteristic frequency (CF). When the CF matches the frequency of a stimulus it is called on-frequency.

Before the IHC-AN synaptic complex, the representation of sound is determined by the movement of the tympanic membrane, the fluctuation of fluids in the inner ear, the vibration of the basilar membrane and the motion of the stereocilia in the inner and outer hair cells. At the IHC-AN complex, however, mechanical energy is converted into an approximately discrete, electrochemical signal that is further processed in the central nervous system. Both the envelope (ENV) information and the temporal fine structure (TFS) information are represented to a certain extent in the timing of the AN discharges (see Section 2.1.1).

The phenomena related to the auditory periphery that are particularly important for the work presented in the current thesis are: (i) mechanical bandpass filtering by the properties of the BM, (ii) amplification provided by the OHCs, (iii) half-wave rectification by the stereocilia, (iv) low-pass filtering by the inertia of the system.

2.1.1. Key characteristics of the periphery signals for this research

The temporal precision of the AN output is especially important for binaural processing if an ITD is present. One important measurement for the temporal precision is the spiking synchrony or phase-locking often expressed as vector strength (VS) (Goldberg and Brown, 1969). Each individual spike is represented as a unit vector with angle α_k corresponding to the spike time within the cycle. The vector strength is defined as

$$v = \left| \frac{1}{K} \sum_{k=1}^{K} \exp(i\alpha_k) \right|, \tag{2.1}$$

with K being the total number of spikes and k indicating the kth spike. If all spikes occur at the same phase of the stimulus waveform VS becomes 1. Conversely, VS approaches 0 if the spike timing is random relative to the phase of the stimulus waveform.

Auditory nerve responses to pure sine tones

The AN fibers fire action potentials according to an underlying von Mises distribution (cf. Fisher, 1993; Ashida *et al.*, 2010; Peterson and Heil, 2020). This distribution follows the phase of the input signal and thus the spike train is phase-locked to the stimulus. It is well known, that phase-locking to the TFS in single fibers of the AN declines above a certain frequency (Rose *et al.*, 1967; Johnson, 1980; Heeringa *et al.*, 2020) (see Figure 2.1) that varies across species (Weiss and Rose, 1988).
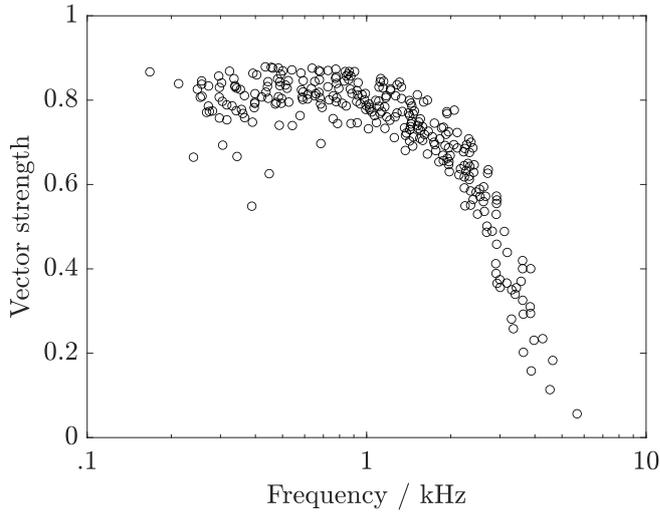
Figure 2.1.: Maximum vector strength for tones at the center frequency of the measured AN fibers as a function of frequency. Data from cat (Johnson, 1980).

One reason of this decline is the low-pass filtering introduced by the inertia of the inner ear (see Section 2.1). The review by Verschooten *et al.* (2019) demonstrated that there is no clear consensus regarding the upper limit of AN phase-locking. In Klug *et al.* (2023) the slope of the VS decline across frequency is described in detail and a lowpass filter is fitted to describe the VS from previous data (see Chapter 4). This VS decline is often posited (e.g. Verschooten *et al.*, 2019) as the explaining factor of the abrupt decrease in pure tone ITD sensitivity (Brughera *et al.*, 2013). However, Klug and Dietz (2022) measured the steepness of the ITD$_{\text{TFS}}$ sensitivity decline across frequency and found it to be much steeper than that of AN VS (see Chapter 5).

Auditory nerve responses to sinusoidally amplitude modulated tones

For research concerning the ITD cues extracted from the stimulus envelope, sinusoidally amplitude modulated (SAM) tones are often used. An example of how to construct such a stimulus is provided in Section 3.4.1 and Eq. (3.6). The spectrum of such a stimulus consists of the carrier frequency f_c and two sidebands at $f_c \pm f_m$, where f_m denotes the modulation frequency (see Figure 2.4 A). The spiking pattern of the AN fibers follow the amplitude modulation of the SAM (an example of this phase-locked response is shown in Figure 3.2). Phase-locking to SAM tones degrades for modulation frequencies between 250 and 500 Hz (see Figure 2.2 B)(Joris and Yin, 1992; Dreyer and Delgutte, 2006). As the modulation frequency increases, the sidebands associated with amplitude modulation are increasingly attenuated by the cochlear filter centered at the CF, thereby reducing the modulation depth of the mechanical input to the hair cells. Thus the VS of the ENV follows a low-pass characteristic similar to that of the TFS, but for a different reason and thus with a lower corner frequency. Unlike the TFS, the VS to ENV is level-dependent in an non-monotonic way: The phase-locking to the SAM tone increases to peak value at around 20 dB and then decreases with increasing level (see Figure 2.2 A)(Joris and Yin, 1992; Dreyer and Delgutte, 2006).
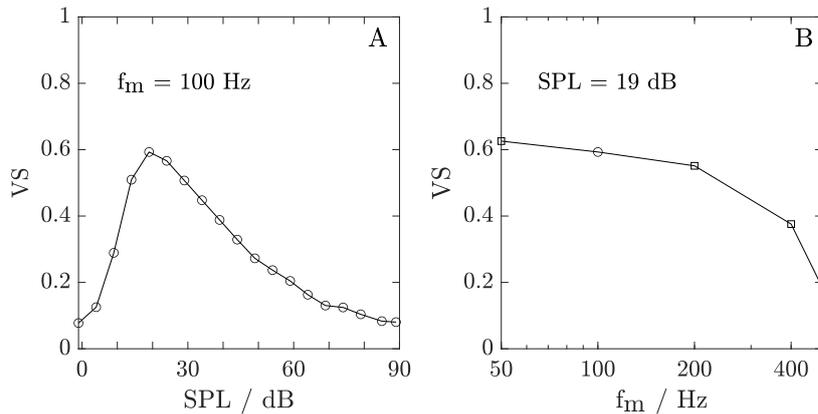

Figure 2.2.: **A** Synchrony of an AN fiber to SAM tones expressed as vector strength across sound pressure level. **B** Vector strength across modulation frequency f_m for the same fiber. On frequency recording: $f_c = CF = 27.2\,kHz$ (Joris and Yin, 1992, Fig. 9).

2.1.2. Models of the periphery

Computational modeling of the auditory periphery has been an active field for the past decades (Lopez-Poveda and Meddis, 2001; Heinz *et al.*, 2001; Hohmann, 2002; Meddis *et al.*, 2010; Heil *et al.*, 2011; Verhulst *et al.*, 2012; Zilany *et al.*, 2014). Here, the focus is on the system-theoretical aspects of modeling auditory processing of tones and SAM tones. For a more detailed presentation and comparison of several models, the reader is guided to Osses Vecchi *et al.* (2022). Here is just a brief overview of the different varieties of models.

Most models of the auditory periphery can be divided into three categories: Biophysical, phenomenological and functional-effective models. Biophysical models (e.g., Verhulst *et al.*, 2015) are designed with a high level of anatomical detail (e.g. transmission line models) to explore how system properties emerge from biological mechanisms. Phenomenological models (e.g., Zilany *et al.*, 2014; Bruce *et al.*, 2018, as used in Chapter 3) are models that primarily predict physiological properties (e.g. spiking rate, peristimulus time histogram) of the system using an abstract processing principle without a direct physiological interpretation. Functional-effective models (e.g., Dau *et al.*, 1996; Hohmann, 2002), simulate the input-output behavior of the auditory system as observed in psychoacoustic experiments, but they do not explicitly model the exact internal physical mechanisms involved in this process.
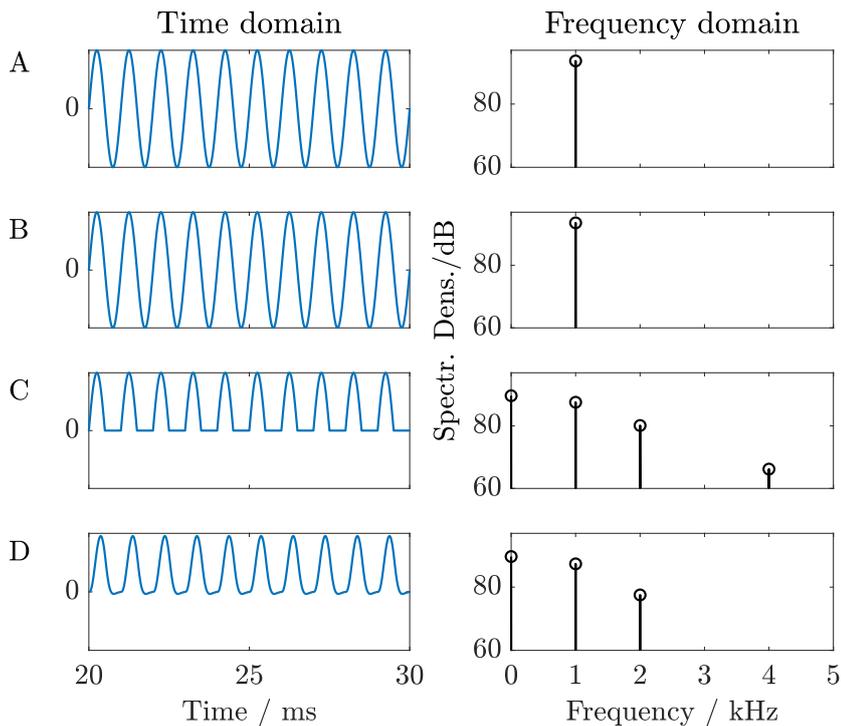
Figure 2.3.: The time domain signal (left) and its spectrum (right) in the on-frequency channel after different processing stages. **A** The signal is a pure tone at 1 kHz. **B** Signal after band-pass filtering: no effect on the pure tone. **C** Signal after half-wave rectification: 0-Hz offset component and higher harmonics are introduced. **D** Signal after low-pass filter: higher harmonics are damped.

Functional-effective models are very successful in reproducing the signal properties of the periphery (cf. Section 2.1.1). The signals relevant for this work, the spectra before and after band-pass filtering, half-wave rectification, and low-pass filtering are briefly shown here for tones (Figure 2.3) and for SAM tones (Figure 2.4). The physiology of the auditory mechanisms involved with these processes is presented in Section 2.1.
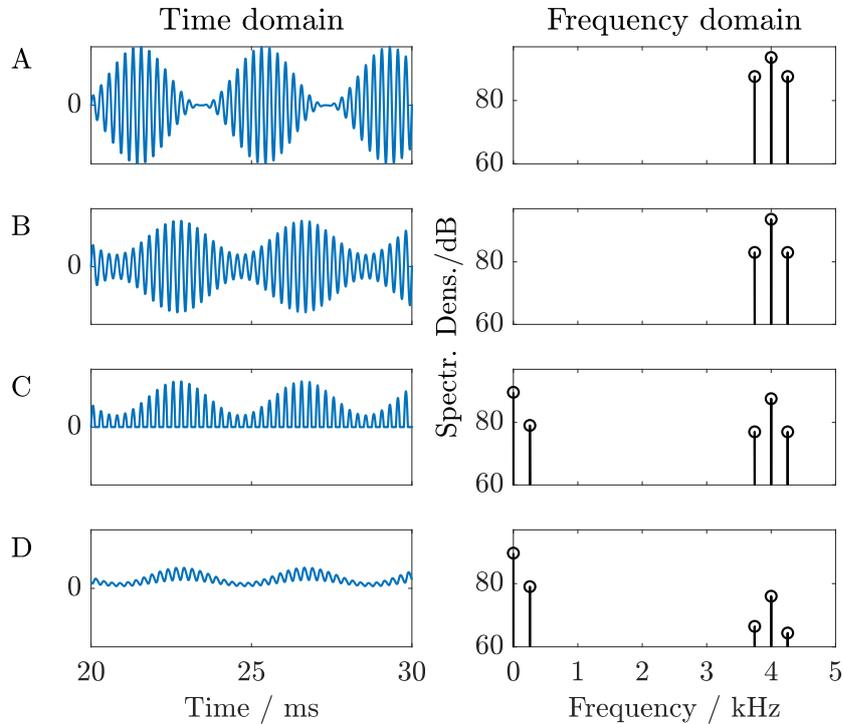
Figure 2.4.: The time domain signal (left) and its spectrum in the on-frequency channel after different processing stages. **A** The signal is an SAM tone with a 4-kHz carrier frequency and 128-Hz modulation frequency. **B** Signal after band-pass filtering: the side bands are attenuated. **C** Signal after half-wave rectification: 0-Hz offset and envelope component (higher harmonics not shown) are introduced. **D** Signal after low-pass filter: higher harmonics are damped.

2.2. First binaural interaction: Processing in the superior olivary complex

This section is about the superior olivary complex (SOC), the anatomical structure of first binaural interaction in the brainstem. Understanding how binaural information is extracted from the signals arriving from the two ears involves understanding how the SOC receives information from the AN (see Figure 2.5). The cochlear nucleus (CN) delivers the information from the AN into the SOC. The spherical bushy cells (SBCs) of the anteroventral cochlear nucleus (AVCN) project excitatory inputs into the SOC. The globular bushy cells (GBCs) project excitatory inputs into the contralateral medial nucleus of the trapezoid body (MNTB) which in return sends inhibitory inputs into the SOC. Both SBCs and GBCs encode timing and intensity information of arriving sounds. For details, see Osen (1969); Kil *et al.* (1995); Grothe *et al.* (2010).
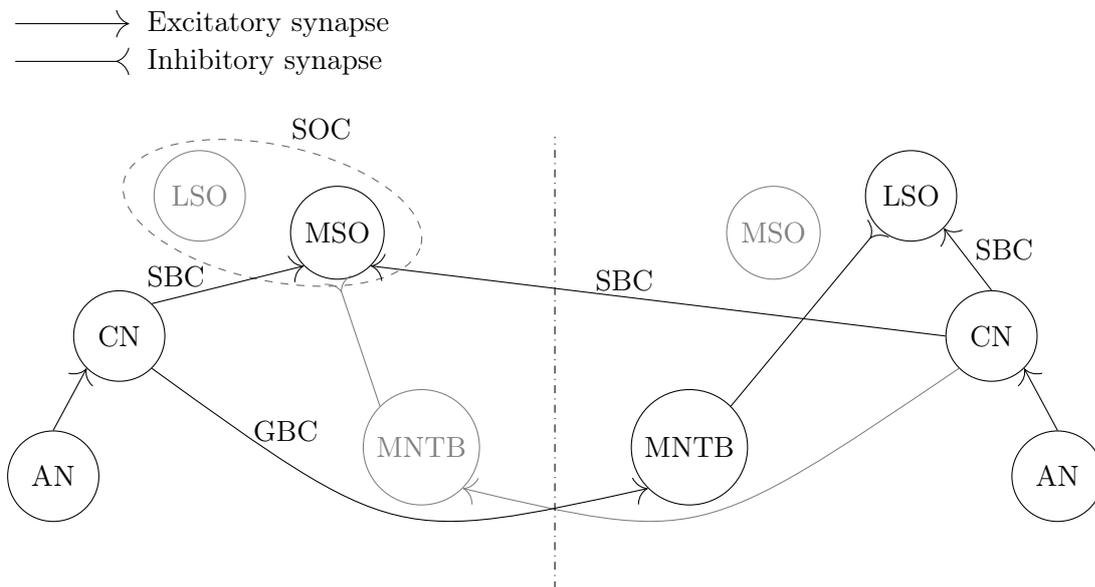
Figure 2.5.: Schematic of the binaural pathway in the brainstem.

The SOC can be divided into different areas or principal nuclei, the medial superior olive (MSO) and the lateral superior olive (LSO). Sections 2.2.1 & 2.2.2 below provide further information on both of these nuclei. Important properties are the type of binaural interaction and the time constants of the individual neurons. The neurons found in the MSO are predominantly of the excitatory-excitatory (EE) type. These neurons display rapid intrinsic electrical resonances and low input impedances, making them well-suited for coincidence detection and, consequently, for processing ITD cues in the TFS of sound waves (Remme *et al.*, 2014). Conversely, neurons in the LSO are predominantly of the excitatory-inhibitory (EI) processing principle and exhibit low-pass electrical properties, indicating their greater efficiency in extracting information from the slower modulations in the amplitude envelopes of sound waves (Remme *et al.*, 2014). These findings, illustrated in Figure 2.6, suggest that different regions or neurons of the SOC are specialized to process different aspects of interaural differences.
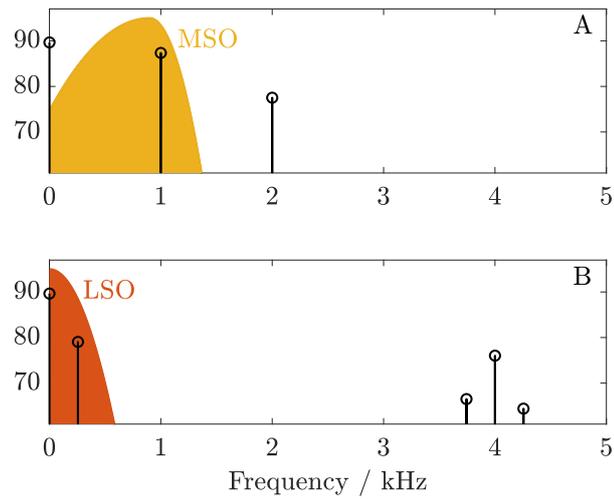
Figure 2.6.: **A** Conceptual representation of MSO sensitivity in the spectrum of a pure tone in the peripheral on-frequency signal. Suitable for encoding TFS information. **B** Conceptual representation of LSO sensitivity in the spectrum of an SAM tone in the peripheral on-frequency signal. Suitable for encoding ENV and ILD information.

2.2.1. Processing in the medial superior olive

The medial superior olive (MSO) is one of several auditory brainstem nuclei that provides the first level of processing for sound localization cues. The principal neurons of the MSO are capable of extracting short ITDs from converging binaural inputs (Cant and Casseday, 1986; Smith *et al.*, 1993; Grothe *et al.*, 2010). To detect these rapid time-varying cues, time difference encoding auditory neurons in the MSO exhibit biophysical specializations, such as low-voltage-activated potassium currents, that allow them to signal with high temporal fidelity (Svirskis *et al.*, 2004). The MSO neurons receive excitatory input from the SBCs from both hemispheres (Stotler, 1953). Despite the fact that the MSO also receives inhibitory inputs (Kuwabara and Zook, 1992; Cant and Hyson, 1992) it is commonly accepted that the MSO output is dominated by the excitatory-excitatory (EE) coincidence detection (Colburn *et al.*, 1990).

At low stimulus frequencies, the MSO output rate modulates with the ITD_{TFS} of a stimulus (Goldberg and Brown, 1969; Yin and Chan, 1990; Grothe *et al.*, 2010). The Rate-IPD functions are generally sinusoidal-like in shape, sometimes modeled as von Mises distribution (Jörg Encke), and their amplitude is symmetrically damped either side of their tuning maximum (see Figure 2.7).
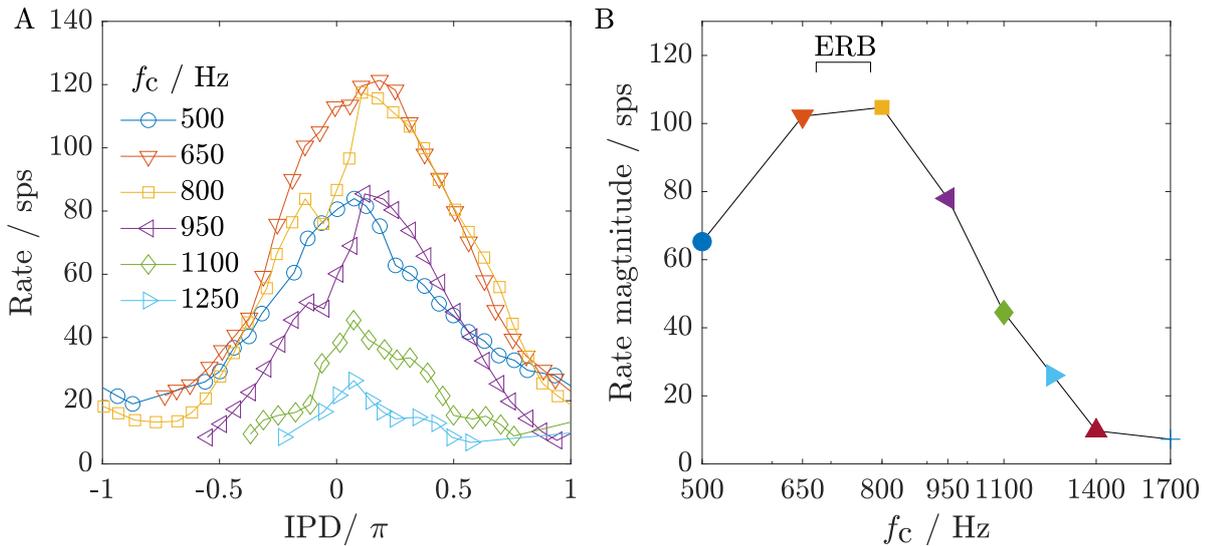
Figure 2.7.: Binaural responses of an MSO cell to changes in the frequency of stimulation (f_c) from Yin and Chan (1990, Fig. 10). **A** Rate-IPD functions for frequencies that cover the response range of the cell. **B** Magnitude of the Rate-IPD function (difference between lowest and highest rate value) across f_c. The black bracket marks the equivalent rectangular bandwidth (ERB) (Glasberg and Moore, 1990) of the peripheral filter according to the CF.

2.2.2. Processing in the lateral superior olive

The lateral superior olive (LSO) is another nucleus where binaural neurons receive inputs originating from the two ears and encode information relevant for sound localization (Tollin, 2003). The principal neurons of the LSO receive excitatory input from the ipsilateral SBCs (Cant and Casseday, 1986; Schwartz, 1992; Owrutsky *et al.*, 2021) and inhibitory input from the contralateral GBCS via the MNTB (Spangler *et al.*, 1985; Sanes, 1990; Owrutsky *et al.*, 2021). This EI circuit can be understood as a subtraction of the signals received from the two ears, which determines the specialization to encode ILDs. Physiological recordings show that the response pattern of LSO neurons depend on the ILD (see Figure 2.8 A; Boudreau and Tsuchitani, 1968; Joris and Yin, 1995; Tollin and Yin, 2002; Tsai *et al.*, 2010). The spike rate of an LSO neuron becomes low when the more intense sound input is on the contralateral ear and high when it is on the ipsilateral ear. Furthermore, the LSO rate changes periodically with the ITD_{ENV} of AM sounds (see Figure 2.8 B; Joris and Yin, 1995; Joris, 1996; Joris and Yin, 1998; Tollin, 2003; Tsai *et al.*, 2010). The troughs of the Rate-ITD_{ENV} functions measured in response to binaural AM stimuli at different modulation frequencies (f_m) commonly align to the same ITD_{ENV} value across different values of f_m (see Figure 2.8 C, Joris and Yin (1995)).
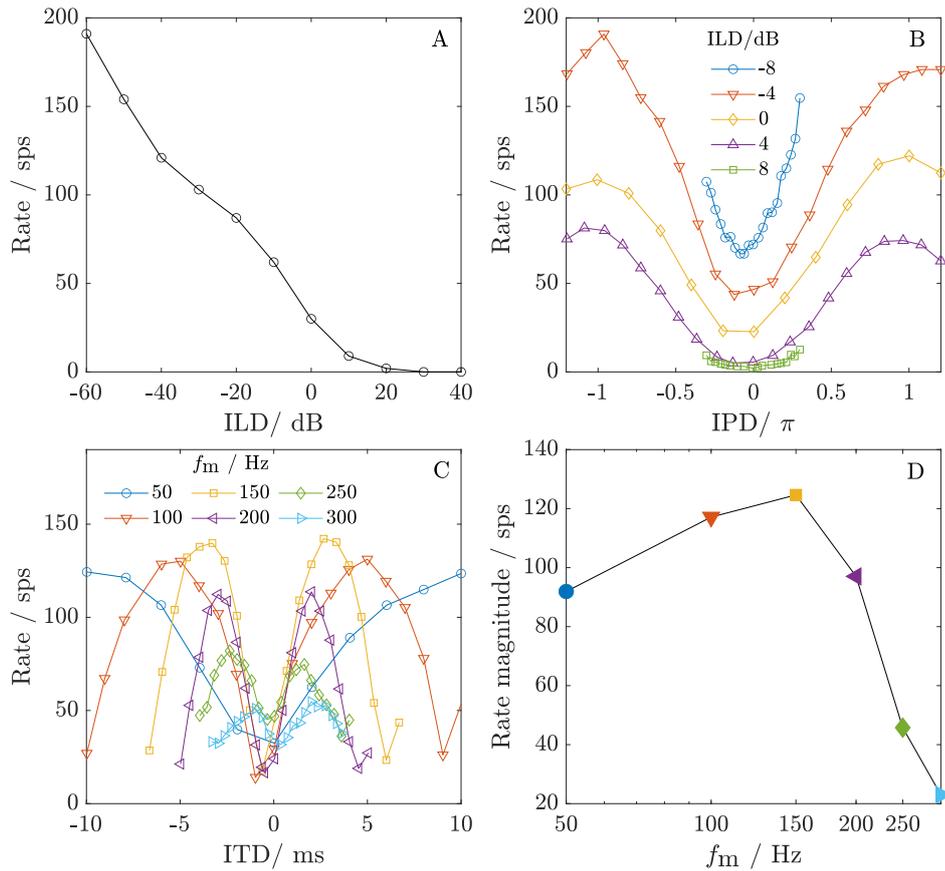
Figure 2.8.: Binaural responses of LSO cells. **A** Rate ILD function from an LSO cell (Joris and Yin, 1995, Fig. 9 A). **B** Rate IPD_{ENV} functions for five different ILDs (Joris and Yin, 1995, Fig. 8). **C** Rate ITD_{ENV} functions for six different modulation frequencies that cover the response range of the cell (CF = 12 kHz)(Joris and Yin, 1995, Fig. 11 B). **D** Magnitude of the Rate-IPD_{ENV} function in C (difference between lowest and highest rate value) across f_m.

2.3. References

Ashida, G., Wagner, H., and Carr, C. E. (**2010**). "Processing of phase-locked spikes and periodic signals," in *Analysis of Parallel Spike Trains*, edited by S. Grün and S. Rotter (Springer US, Boston, MA), pp. 59–74, doi: `10.1007/978-1-4419-5675-0_4`.

Ashmore, J. F. (**1987**). "A fast motile response in guinea-pig outer hair cells: the cellular basis of the cochlear amplifier.," The Journal of Physiology **388**(1), 323–347, doi: `10.1113/jphysiol.1987.sp016617`.

Boudreau, J. C., and Tsuchitani, C. (**1968**). "Binaural interaction in the cat superior olive s segment.," Journal of Neurophysiology **31**(3), 442–454, doi: `10.1152/jn.1968.31.3.442` pMID: 5687764.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001a**). "Binaural processing model based on contralateral inhibition. i. model structure," J. Acoust. Soc. Am. **110**(2), 1074–1088, doi: `10.1121/1.1383297`.

Bruce, I. C., Erfani, Y., and Zilany, M. S. (**2018**). "A phenomenological model of the synapse between the inner hair cell and auditory nerve: Implications of limited neurotransmitter release sites," Hearing Research **360**, 40–54, doi: `10.1016/j.heares.2017.12.016` computational models of the auditory system.

Brughera, A., Dunai, L., and Hartmann, W. M. (**2013**). "Human interaural time difference thresholds for sine tones: The high-frequency limit," J. Acoust. Soc. Am. **133**(5), 2839–2855, doi: `10.1121/1.4795778`.

Cant, N. B., and Casseday, J. H. (**1986**). "Projections from the anteroventral cochlear nucleus to the lateral and medial superior olivary nuclei," Journal of Comparative Neurology **247**(4), 457–476, doi: `10.1002/cne.902470406`.

Cant, N. B., and Hyson, R. L. (**1992**). "Projections from the lateral nucleus of the trapezoid body to the medial superior olivary nucleus in the gerbil," Hearing research **58**(1), 26–34, doi: `10.1016/0378-5955(92)90005-8`.

Colburn, H. S., Yan-an, H., and Culotta, C. P. (**1990**). "Coincidence model of mso responses," Hearing research **49**(1), 335–346, doi: `10.1016/0378-5955(90)90112-3`.

Dau, T., Püschel, D., and Kohlrausch, A. (**1996**). "A quantitative model of the "effective" signal processing in the auditory system. i. model structure," The Journal of the Acoustical Society of America **99**(6), 3615–3622, doi: `10.1121/1.414959`.

Dreyer, A., and Delgutte, B. (**2006**). "Phase locking of auditory-nerve fibers to the envelopes of high-frequency sounds: Implications for sound localization," Journal of Neurophysiology **96**(5), 2327–2341, doi: `10.1152/jn.00326.2006` pMID: 16807349.

Eurich, B., Encke, J., Ewert, S. D., and Dietz, M. (**2022**). "Lower interaural coherence in off-signal bands impairs binaural detection," The Journal of the Acoustical Society of America **151**(6), 3927–3936, doi: `10.1121/10.0011673`.

Fisher, N. I. (**1993**). *Statistical Analysis of Circular Data* (Cambridge University Press).

Glasberg, B. R., and Moore, B. C. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hearing Research **47**(1), 103–138, doi: `10.1016/0378-5955(90)90170-T`.

Goldberg, J. M., and Brown, P. B. (**1969**). "Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization.," Journal of Neurophysiology **32**(4), 613–636, doi: `10.1152/jn.1969.32.4.613`.

Greenwood, D. D. (**1961**). "Critical bandwidth and the frequency coordinates of the basilar membrane," The Journal of the Acoustical Society of America **33**(10), 1344–1356, doi: `10.1121/1.1908437`.

Grothe, B., Pecka, M., and McAlpine, D. (**2010**). "Mechanisms of sound localization in mammals," Physiological Reviews **90**(3), 983–1012, doi: `10.1152/physrev.00026.2009`.

Heeringa, A. N., Zhang, L., Ashida, G., Beutelmann, R., Steenken, F., and Köppl, C. (**2020**). "Temporal coding of single auditory nerve fibers is not degraded in aging gerbils," Journal of Neuroscience **40**(2), 343–354, doi: `10.1523/JNEUROSCI.2784-18.2019`.

Heil, P., Neubauer, H., and Irvine, D. R. F. (**2011**). "An improved model for the rate–level functions of auditory-nerve fibers," Journal of Neuroscience **31**(43), 15424–15437, doi: `10.1523/JNEUROSCI.1638-11.2011`.

Heinz, M. G., Colburn, H. S., and Carney, L. H. (**2001**). "Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the Auditory Nerve," Neural Computation **13**(10), 2273–2316, doi: `10.1162/089976601750541804`.

Hohmann, V. (**2002**). "Frequency analysis and synthesis using a gammatone filterbank," Acta Acustica united with Acustica **88**(3), 433–442.

Hudspeth, A. J. (**1985**). "The cellular basis of hearing: The biophysics of hair cells," Science **230**(4727), 745–752, doi: `10.1126/science.2414845`.

Johnson, D. H. (**1980**). "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," The Journal of the Acoustical Society of America **68**(4), 1115–1122, doi: `10.1121/1.384982`.

Joris, P. X. (**1996**). "Envelope coding in the lateral superior olive. ii. characteristic delays and comparison with responses in the medial superior olive," Journal of Neurophysiology **76**(4), 2137–2156, doi: `10.1152/jn.1996.76.4.2137`.

Joris, P. X., and Yin, T. C. (**1992**). "Responses to amplitude-modulated tones in the auditory nerve of the cat," The Journal of the Acoustical Society of America **91**(1), 215–232, doi: `10.1121/1.402757`.

Joris, P. X., and Yin, T. C. (**1995**). "Envelope coding in the lateral superior olive. I. Sensitivity to interaural time differences," Journal of Neurophysiology **73**(3), 1043–1062, doi: `10.1152/jn.1995.73.3.1043`.

Joris, P. X., and Yin, T. C. T. (**1998**). "Envelope coding in the lateral superior olive. iii. comparison with afferent pathways," Journal of Neurophysiology **79**(1), 253–269, doi: `10.1152/jn.1998.79.1.253`.

2. Processing in the auditory pathway

Jürgens, T., and Brand, T. (**2009**). "Microscopic prediction of speech recognition for listeners with normal hearing in noise using an auditory model," The Journal of the Acoustical Society of America **126**(5), 2635–2648, doi: `10.1121/1.3224721`.

Kil, J., Hkageyama, G., Semple, M. N., and Kitzes, L. M. (**1995**). "Development of ventral cochlear nucleus projections to the superior olivary complex in gerbil," Journal of Comparative Neurology **353**(3), 317–340, doi: `10.1002/cne.903530302`.

Klein-Hennig, M., Dietz, M., Hohmann, V., and Ewert, S. D. (**2011**). "The influence of different segments of the ongoing envelope on sensitivity to interaural time delays," The Journal of the Acoustical Society of America **129**(6), 3856–3872, doi: `10.1121/1.3585847`.

Klug, J., and Dietz, M. (**2022**). "Frequency dependence of sensitivity to interaural phase differences in pure tones," The Journal of the Acoustical Society of America **152**(6), 3130–3141, doi: `10.1121/10.0015246`.

Klug, J., Encke, J., and Dietz, M. (**2023**). "Characterization of the decline in the auditory nerve phase locking at high frequencies," JASA Express Letters **3**(7), 074403, doi: `10.1121/10.0020267`.

Kuwabara, N., and Zook, J. M. (**1992**). "Projections to the medial superior olive from the medial and lateral nuclei of the trapezoid body in rodents and bats," Journal of Comparative Neurology **324**(4), 522–538, doi: `10.1002/cne.903240406`.

Lopez-Poveda, E. A., and Meddis, R. (**2001**). "A human nonlinear cochlear filterbank," The Journal of the Acoustical Society of America **110**(6), 3107–3118, doi: `10.1121/1.1416197`.

Meddis, R., Lopez-Poveda, E. A., Fay, R. R., and Popper, A. N. (**2010**). *Computational Models of the Auditory System* (Springer, New York).

Osen, K. K. (**1969**). "The intrinsic organization of the cochlear nuclei in the cat," Acta Oto-Laryngologica **67**(2-6), 352–359, doi: `10.3109/00016486909125462` pMID: 5374653.

Osses Vecchi, A., Varnet, L., Carney, L. H., Dau, T., Bruce, I. C., Verhulst, S., and Majdak, P. (**2022**). "A comparative study of eight human auditory models of monaural processing," Acta Acust. **6**, 17, doi: `10.1051/aacus/2022008`.

Owrutsky, Z. L., Benichoux, V., and Tollin, D. J. (**2021**). "Binaural hearing by the mammalian auditory brainstem: Joint coding of interaural level and time differences by the lateral superior olive," in *Binaural Hearing*, edited by R. Y. Litovsky, M. J. Goupell, R. R. Fay, and A. N. Popper (Springer International Publishing, Cham), pp. 113–144, doi: `10.1007/978-3-030-57100-9_5`.

Peterson, A. J., and Heil, P. (**2020**). "Phase locking of auditory nerve fibers: The role of lowpass filtering by hair cells," Journal of Neuroscience **40**(24), 4700–4714, doi: `10.1523/JNEUROSCI.2269-19.2020`.

Plack, C. (**2013**). *The sense of hearing: Second Edition* (Psychology Press, New York).

Remme, M. W. H., Donato, R., Mikiel-Hunter, J., Ballestero, J. A., Foster, S., Rinzel, J., and McAlpine, D. (**2014**). "Subthreshold resonance properties contribute to the efficient coding of auditory spatial cues," Proceedings of the National Academy of Sciences **111**(22), E2339–E2348, doi: `10.1073/pnas.1316216111`.

Robles, L., and Ruggero, M. A. (**2001**). "Mechanics of the mammalian cochlea," Physiological Reviews **81**(3), 1305–1352, doi: `10.1152/physrev.2001.81.3.1305`.

Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (**1967**). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey.," Journal of Neurophysiology **30**(4), 769–793, doi: `10.1152/jn.1967.30.4.769`.

Sanes, D. (**1990**). "An in vitro analysis of sound localization mechanisms in the gerbil lateral superior olive," Journal of Neuroscience **10**(11), 3494–3506, doi: `10.1523/JNEUROSCI.10-11-03494.1990`.

Schwartz, I. R. (**1992**). "The superior olivary complex and lateral lemniscal nuclei," in *The Mammalian Auditory Pathway: Neuroanatomy*, edited by D. B. Webster, A. N. Popper, and R. R. Fay (Springer New York, New York, NY), pp. 117–167, doi: `10.1007/978-1-4612-4416-5_4`.

Smith, P. H., Joris, P. X., and Yin, T. C. T. (**1993**). "Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: Evidence for delay lines to the medial superior olive," Journal of Comparative Neurology **331**(2), 245–260, doi: `10.1002/cne.903310208`.

Søndergaard, P. L., and Majdak, P. (**2013**). "The auditory modeling toolbox," in *The Technology of Binaural Listening*, edited by J. Blauert (Springer Berlin Heidelberg, Berlin, Heidelberg), pp. 33–56, `https://www.amtoolbox.org/`.

Spangler, K. M., Warr, W. B., and Henkel, C. K. (**1985**). "The projections of principal cells of the medial nucleus of the trapezoid body in the cat," Journal of Comparative Neurology **238**(3), 249–262, doi: `h10.1002/cne.902380302`.

Stotler, W. A. (**1953**). "An experimental study of the cells and connections of the superior olivary complex of the cat," Journal of Comparative Neurology **98**(3), 401–431, doi: `10.1002/cne.900980303`.

Svirskis, G., Kotak, V., Sanes, D. H., and Rinzel, J. (**2004**). "Sodium along with low-threshold potassium currents enhance coincidence detection of subthreshold noisy signals in mso neurons," Journal of Neurophysiology **91**(6), 2465–2473, doi: `10.1152/jn.00717.2003` pMID: 14749317.

Tollin, D. J. (**2003**). "The lateral superior olive: A functional role in sound source localization," The Neuroscientist **9**(2), 127–143, doi: `10.1177/1073858403252228` pMID: 12708617.

Tollin, D. J., and Yin, T. C. T. (**2002**). "The coding of spatial location by single units in the lateral superior olive of the cat. I. Spatial receptive fields in azimuth," Journal of Neuroscience **22**(4), 1454–1467, doi: `10.1523/JNEUROSCI.22-04-01454.2002`.

Tsai, J. J., Koka, K., and Tollin, D. J. (**2010**). "Varying overall sound intensity to the two ears impacts interaural level difference discrimination thresholds by single neurons in the lateral superior olive," Journal of Neurophysiology **103**(2), 875–886, doi: `10.1152/jn.00911.2009`.

Verhulst, S., Bharadwaj, H. M., Mehraei, G., Shera, C. A., and Shinn-Cunningham, B. G. (**2015**). "Functional modeling of the human auditory brainstem response to broadband stimulation," The Journal of the Acoustical Society of America **138**(3), 1637–1659, doi: `10.1121/1.4928305`.

Verhulst, S., Dau, T., and Shera, C. A. (**2012**). "Nonlinear time-domain cochlear model for transient stimulation and human otoacoustic emission," The Journal of the Acoustical Society of America **132**(6), 3842–3848, doi: `10.1121/1.4763989`.

Verschooten, E., Shamma, S., Oxenham, A., Moore, B., Joris, P., Heinz, M., and Plack, C. (**2019**). "The upper frequency limit for the use of phase locking to code temporal fine structure in humans: A compilation of viewpoints," Hearing Research **377**, doi: `10.1016/j.heares.2019.03.011`.

von Békésy, G. (**1970**). "Travelling waves as frequency analysers in the cochlea," Nature **225**, 1207–1209, doi: `10.1038/2251207a0`.

Weiss, T., and Rose, C. (**1988**). "A comparison of synchronization filters in different auditory receptor organs," Hearing Research **33**(2), 175–179, doi: `10.1016/0378-5955(88)90030-5`.

Yin, T. C., and Chan, J. C. (**1990**). "Interaural time sensitivity in medial superior olive of cat," Journal of Neurophysiology **64**(2), 465–488, doi: `10.1152/jn.1990.64.2.465`.

Zilany, M. S. A., Bruce, I. C., and Carney, L. H. (**2014**). "Updated parameters and expanded simulation options for a model of the auditory periphery," The Journal of the Acoustical Society of America **135**(1), 283–286, doi: `10.1121/1.4837815`.

3. Neural rate difference model can account for lateralization of high-frequency stimuli

3.1. Abstract

Lateralization of complex high-frequency sounds is conveyed by interaural level differences (ILDs) and interaural time differences (ITDs) in the envelope. In this work, we constructed an auditory model and simulated data from three previous behavioral studies obtained with, in total, over 1000 different amplitude-modulated stimuli. We combined a well-established auditory periphery model with a functional count-comparison model for binaural excitatory-inhibitory (EI) interaction. After parameter optimization of the EI-model stage, the hemispheric rate-difference between pairs of EI-model neurons relates linearly to the extent of laterality in human listeners. If a certain ILD and a certain envelope ITD each cause a similar extent of laterality, they also produce a similar rate difference in the same model neurons. After parameter optimization, the model accounts for $95.7\,\%$ of the variance in the largest dataset, in which amplitude modulation depth, rate of modulation, modulation exponent, ILD, and envelope ITD were varied. The model also accounts for $83\,\%$ of the variances in each of the other two data sets using the same EI-model parameters.

----

3. Neural rate difference model can account for lateralization of high-frequency stimuli

3.2. Introduction

Accurate sound localization requires precise neural mechanisms for processing relevant binaural cues, such as interaural time difference (ITD, here denoted by Δt) and interaural level difference (ILD). The first stage of neural integration of binaural information is located in the superior olivary complex (SOC) in the auditory brainstem, where projections from the left and the right side converge (for review see Grothe *et al.*, 2010). More specifically, the medial superior olive (MSO) plays a dominant role in encoding fine-structure ITDs of low-frequency sounds, while neurons in the lateral superior olive (LSO) are often sensitive to both ILDs and ITDs, including envelope ITDs (for review see Tollin, 2003). A fundamental question in the study of binaural-information processing is how to relate the neuronal representations of these cues to the evoked percepts.

Historically, the mechanism for ITD encoding was envisaged to be formed by an array of binaural coincidence-detecting neurons receiving differently delayed inputs from the left and right ear. These internal delays were thought to compensate for the respective external ITD (Jeffress, 1948). Later anatomical and physiological studies found such circuitry in the auditory brainstem of birds (Carr and Konishi, 1990; Köppl and Carr, 2008). Based on this delay-line approach, binaural perception has most often been modeled and explained by interaural cross-correlation. Such a coincidence-detecting model unit responds maximally when the relative internal delay between its bilateral inputs compensates for the external ITD (e.g., Jeffress, 1948; Lindemann, 1986; Bernstein and Trahiotis, 2003; Stern and Shear, 1996; Colburn, 1977). However, a delay-line mechanism may not be operational in the mammalian binaural pathway (e.g., Grothe *et al.*, 2010), but the issue is still under debate (Leibold and Grothe, 2015; Yin *et al.*, 2019; Joris and van der Heijden, 2019).

For the processing of ITDs in the envelope of high-frequency stimuli, there is a more fundamental discrepancy between animal physiological studies and the most comprehensive models of human perception: Most neurons sensitive to envelope ITDs, especially in the LSO, receive excitatory input originating from the ipsilateral side and inhibitory input from the contralateral side (Tollin, 2003), but models that account for envelope ITD-based lateralization or discrimination commonly use an excitatory-excitatory or multiplicative interaction (e.g., Bernstein and Trahiotis, 2003, 2012). Despite this discrepancy, these models can account for most perceptual data with high accuracy. Notably, Bernstein and Trahiotis (2012) were able to account for 94 % of the variance of their psychoacoustic data obtained using 960 different stimuli that were formed by varying five different stimulus parameters: amplitude modulation depth, rate of modulation, modulation exponent, ILD, and envelope ITD.

The primary goal of the current study is to investigate whether a model framework that employs neither a delay-line scheme nor a multiplication-based cross-correlation can still account for the psychoacoustic data of Bernstein and Trahiotis (2003, 2012) and of Dietz *et al.* (2015). We demonstrate here that the rate difference between a left-hemispheric and a right-hemispheric binaural excitatory-inhibitory (EI) model neuron is largely sufficient to explain both envelope-ITD-based and ILD-based lateralization. The EI interaction model reproduces the well-known sigmoidal ILD rate functions comparable to the characteristic response of LSO neurons. For envelope ITDs up to at least 1 ms the rate difference between left and right EI-model neurons increases monotonically with ITD. The monotonic relations are exploited by the decoding stage, mapping the rate difference to a perceptual quantity, the extent of laterality.

3.3. Methods

3.3.1. Model topology
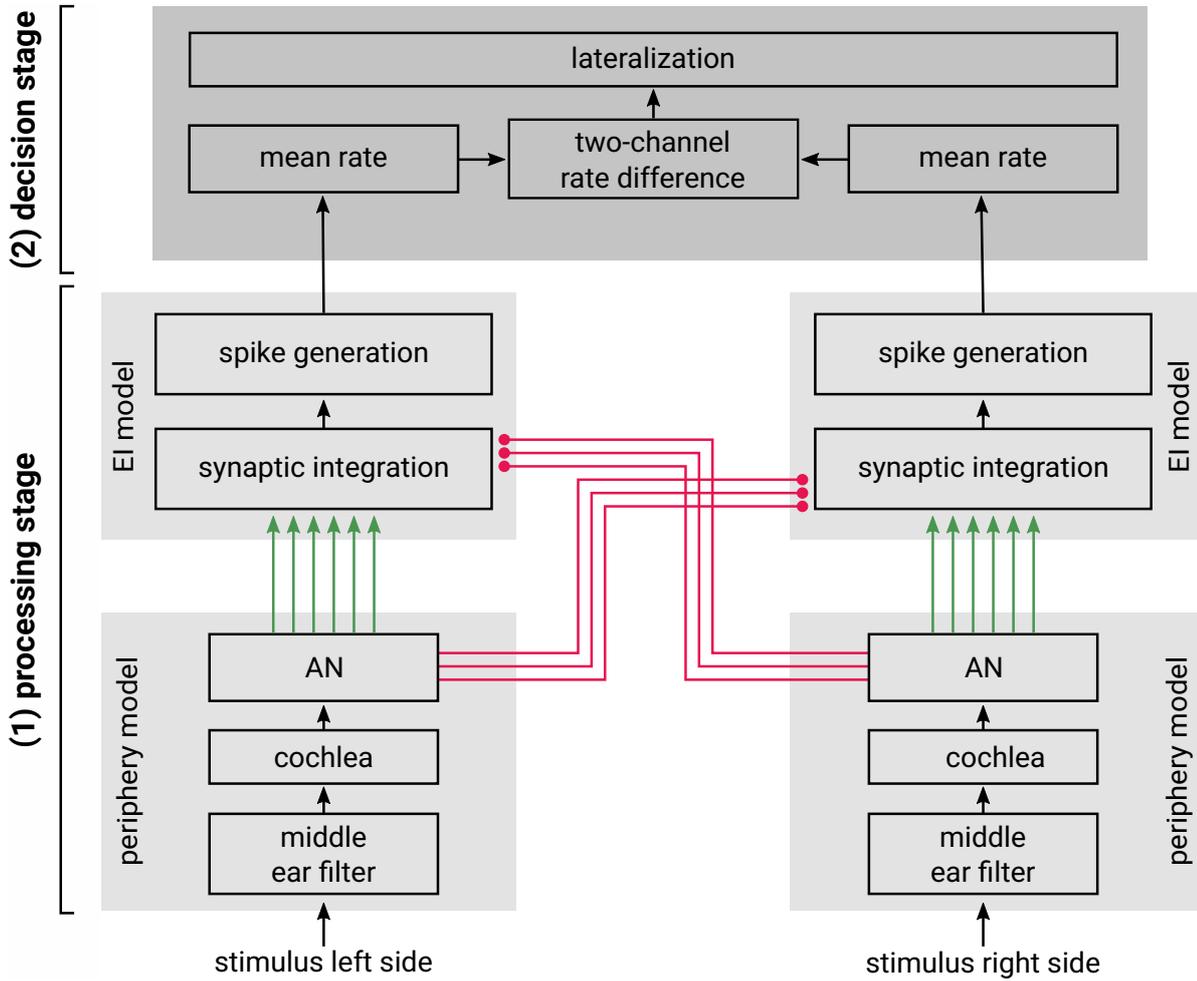
Figure 3.1.: The model structure is subdivided into two parts: (1) The primary processing stage that constitutes both the periphery receiving the binaural sound stimulus as the input, and the excitatory-inhibitory (EI) integration stage that bilaterally receives the excitatory (arrow) and inhibitory (bullet) outputs of the periphery. (2) The decision stage with a simple two-channel rate-difference model that maps to the acoustic pointer (see section 3.3.4) and predicts the extent of laterality.

The physiologically motivated binaural lateralization model (Fig. 3.1) starts with the auditory periphery model of (Bruce *et al.*, 2018). The input to the model is a stimulus in the form of a pressure waveform (Fig. 3.2 A). The stimulus is first processed by a band-pass filter accounting for the processing of the middle ear (Fig. 3.1). Subsequent to the middle-ear filter, the signal is processed by three parallel feed-forward paths: the component 1, the component 2, and the control path (Zilany and Bruce, 2006). The collective response properties of the basilar membrane and the inner hair cells (IHCs) are represented in these pathways accounting for both the passive mechano-electrical transduction at the inner hair cells, as well as the mechano-electrical and electro-mechanical transduction facilitated by the outer hair cells. The filtered signal is converted to receptor potentials of the IHCs (Fig. 3.2 B), where each has its own characteristic frequency (CF). The model also includes a physiologically realistic representation
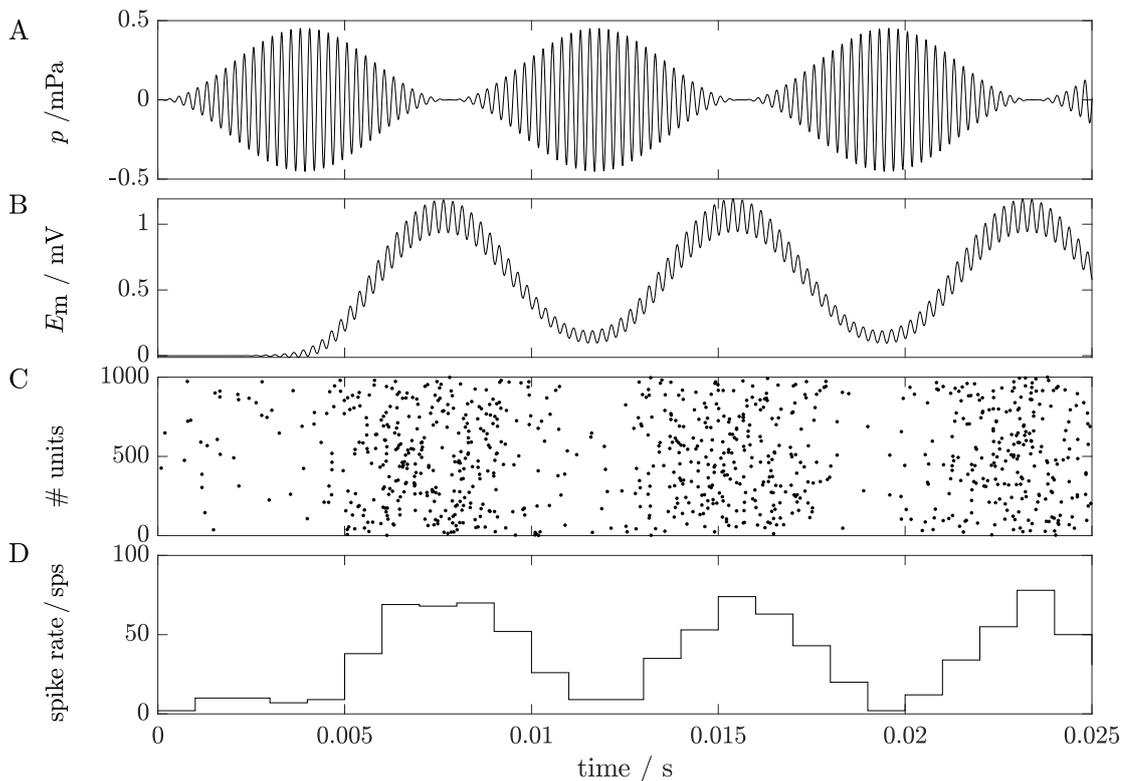
Figure 3.2.: Steady state response of the periphery model (CF = 4 kHz) for a fully modulated, sinusoidally amplitude-modulated (SAM) tone with a modulation frequency f_{m} = 128 Hz and carrier frequency f_{c} = 4 kHz at a stimulus level of 20 dB SPL. **A** Stimulus sound waveform, **B** the IHC receptor potential, **C** the spike raster plot for AN fibers, and **D** the PSTH.

of the synapses between the inner hair cells and the auditory nerve (AN). The output of the model is given by a spike generator that produces a series of AN spikes (Fig. 3.2 C; for a more detailed description of the model see Bruce *et al.*, 2018). Each AN fiber of this model depends on a spontaneous rate r_{s}, on a relative-refractory time r_{rel} and on an absolute-refractory time t_{abs}. In Fig. 3.2 D, the spike phase of the output is plotted in a peristimulus time histogram (PSTH). In a nutshell, the block "periphery model" in Fig. 3.1 transforms the acoustic stimulus into a spiking pattern of AN fibers arranged along the tonotopical axis.

For the binaural interaction stage, we used the coincidence-counting model of Ashida *et al.* (2016). A model neuron of this stage receives excitatory synaptic inputs from ipsilateral AN fibers and inhibitory inputs from contralateral AN fibers (Fig. 3.1), all with the same CF. In the EI-model, two temporal rectangular windows slide along the time axis: for the excitatory fibers from the ipsilateral side and for the inhibitory fibers of the contralateral side. The sum of excitatory spikes, each counting +1, and inhibitory spikes, each counting -δ form an activation variable that can be understood as a surrogate for the membrane potential of a real neuron (Ashida *et al.*, 2017). Once the activation variable reaches a specified response threshold θ, the EI-model neuron generates an output spike. Subsequent to a generated action potential, no further output is possible within the refractory period T, even if synaptic integration during this period may occur (i.e., no zero set after spiking).

3.3.2. Relating EI-model output to experimental LSO data

Characteristic outputs of the EI-stage in response to a $20\,\text{dB}\,\text{SPL}$ sinusoidally amplitude-modulated tone with a carrier frequency (f_c) of $4\,\text{kHz}$ and a modulation frequency (f_m) of $128\,\text{Hz}$ are shown in Figure 3.3. Instead of using spike trains generated by a Poisson process as an input (Ashida *et al.*, 2016), we use a model of the auditory periphery (Bruce *et al.*, 2018) as the front end. The EI-model by (Ashida *et al.*, 2016) was never tested before for such an input and generated only a very sparse output, if any. Therefore, the EI parameters had to be adjusted (see section 3.4.). With these settings, the model produces ILD- and ITD-rate functions (Fig. 3.3 A and 3.3 B) similar to those observed experimentally in the LSO (e.g., Joris and Yin, 1995; Joris, 1996; Joris and Yin, 1998; Tollin and Yin, 2002). As also physiologically measured, the simulated output spike rate of the model varies periodically with ITD (Fig. 3.3 A). The overall response rate increases with ILD, while the shape of the periodic ITD-rate functions remains mostly unaffected by ILD (comparable physiological results can be found in Fig. 8 of Joris and Yin, 1995). For negative values of ILD, which indicate a higher stimulus intensity at the left side, the response rate of the left simulated EI-neuron is higher compared to the right model neuron (Fig. 3.3 B); for positive ILDs, this relation is reversed so that the right model neuron generates more spikes than the left. This effect is comparable to physiological data from Joris and Yin (1995, Fig 9 A).

The ITD rate functions of the left and the right hemisphere are shown in Fig. 3.3 C. The ILD was set to $0\,\text{dB}$ and the response rates of the simulated neurons are shown as a function of ITD (for a physiological comparison, see Fig. 16 B of Joris and Yin, 1998). The trough is not at zero ITD, because the inhibition lasts longer than the excitation. The minimum response is reached when the excitation is centered in the longer inhibition (Ashida *et al.*, 2016), i.e. $\Delta t_\text{worst} = (W_\text{in} - W_\text{ex})/2$. The ITD sensitivity is mirrored at $\Delta t = 0\,\text{ms}$ for the model neurons of the left and right sides. The rate difference between the two hemispheres is shown in Fig. 3.3 D. The function is nearly point symmetric and linear around the coordinate origin. The shape of the rate-difference function depends on the shapes of the left and right ITD rate functions, which, in turn, depend on the stimulus parameters such as ILD, modulation frequency, and modulation depth (see section 3.4.1).
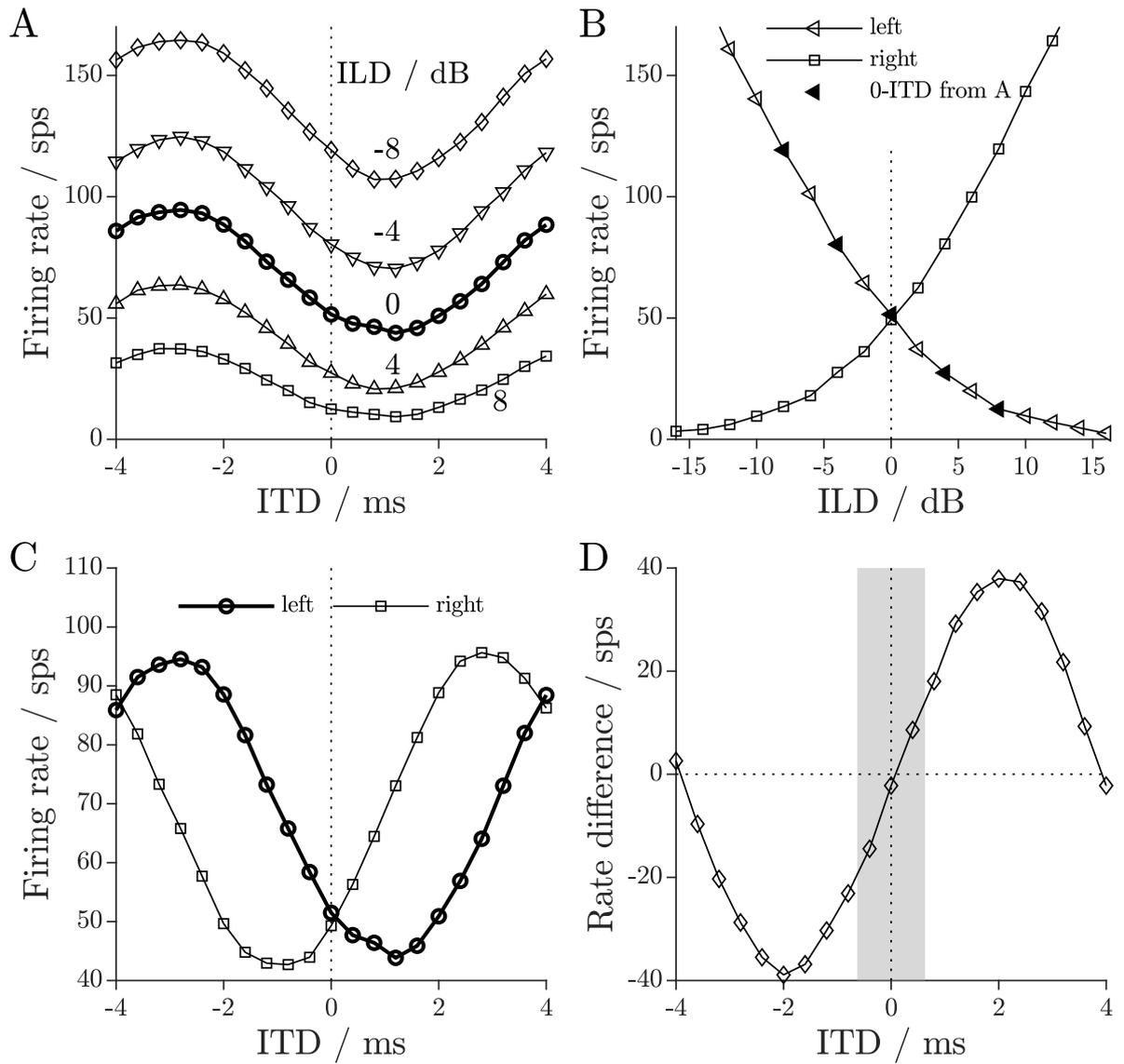
Figure 3.3.: Tuning functions of the EI-stage for ITD and ILD. The model parameters were those derived further below in the Sec. 3.4. (see Table 3.1, best performance). AN input fibers had a CF = 4 kHz. The same stimulus as for Fig. 3.2 was used. (A) ITD rate functions of the EI-model in the left hemisphere for five different ILDs. (B) ILD rate functions for EI-model neurons in the left and right hemisphere (ITD = 0 ms). Filled symbols correspond to data points also shown in panel A. (C) ITD rate functions for the left and right hemisphere (ILD = 0 dB). (D) Rate difference between the left and right EI-model neuron output. The shaded area between ± 0.63 ms corresponds to the approximated physiological range of humans.

3.3.3. Quantifying ITD-information transmission

Figure 3.4 A-D displays the model output of the intermediate processing stages along the tonotopic array for a 4 kHz, sinusoidally amplitude-modulated (SAM) tone. For faithful coding of ITD information, the peripheral stage needs to produce sufficient activity (quantified by spike rate, Fig. 3.4 A) and phase locking (Fig. 3.4 B). The degree of phase locking can be quantified by the vector strength (Goldberg and Brown, 1969). Each individual spike is represented as a unit vector with angle α_k corresponding to the spike time within the cycle. The vector strength is defined as

$$v = \left| \frac{1}{K} \sum_{k=1}^{K} \exp(i\alpha_k) \right|, \tag{3.1}$$

with K being the total number of spikes and k indicating the kth spike. If all spikes occur at a single phase of the stimulus waveform v becomes 1. Phase locking can be visually observed in Figure 3.2 C + D by the synchronized responses of the simulated AN fibers to the envelope of the SAM tone.

Compared to physiological experiments, psychoacoustic measurements of envelope-ITD perception are usually performed at much higher sound levels, e.g., 65 or 75 dB SPL. Since AN fibers with CFs matched to the carrier frequency phase lock very poorly to the envelope at such high levels (Joris and Yin, 1992; Dreyer and Delgutte, 2006), envelope ITDs have to be extracted by other means, presumably by neurons tuned to frequencies different from the signal carrier. Figure 3.4 shows model responses to a higher-level stimulus of 68 dB SPL (compared to 20 dB SPL in Fig. 3.2 and 3.3). It is apparent that off-frequency neurons can encode ITD information, while the response of on-frequency EI-model neurons is generally lower and barely changes with ITD (Fig. 3.4 C).

3.3.4. Decoding the EI response

The output of the AN model along the tonotopic axis (Fig. 3.4 A + B) serves as the input for the central processing stage (Fig. 3.4 C + D), whose output is then used in the decision stage to simulate the extent of laterality. Each EI-model neuron receives a number of excitatory and inhibitory AN model inputs (see Table 3.1) with matching CF. Two EI-model neurons, one from the left and one from the right side (with the same CF) form a pair and the hemispheric rate difference is computed between them:

$$\Delta R_j = R_{jR} - R_{jL} \tag{3.2}$$

with R_{jR} being the rate of an EI-model-neuron in the right hemisphere at the respective CF(j). The mean spike rate difference between the left and right hemisphere is

$$\Delta\overline{R} = \frac{\sum_j \Delta R_j}{N}, \tag{3.3}$$

with N being the number of frequency channels. The mean rate difference $\Delta\overline{R}$ is later converted into the perceived extent of laterality measured in the psychoacoustic experiments [Eq. (3.4)]. In previous studies, various central read-out mechanisms were used (for a review, see Dietz *et al.*, 2018). Kelvasa and Dietz (2015) showed that the hemispheric response difference, averaged across the tonotopic array of LSO model neurons, is proportional to azimuthal sound-source
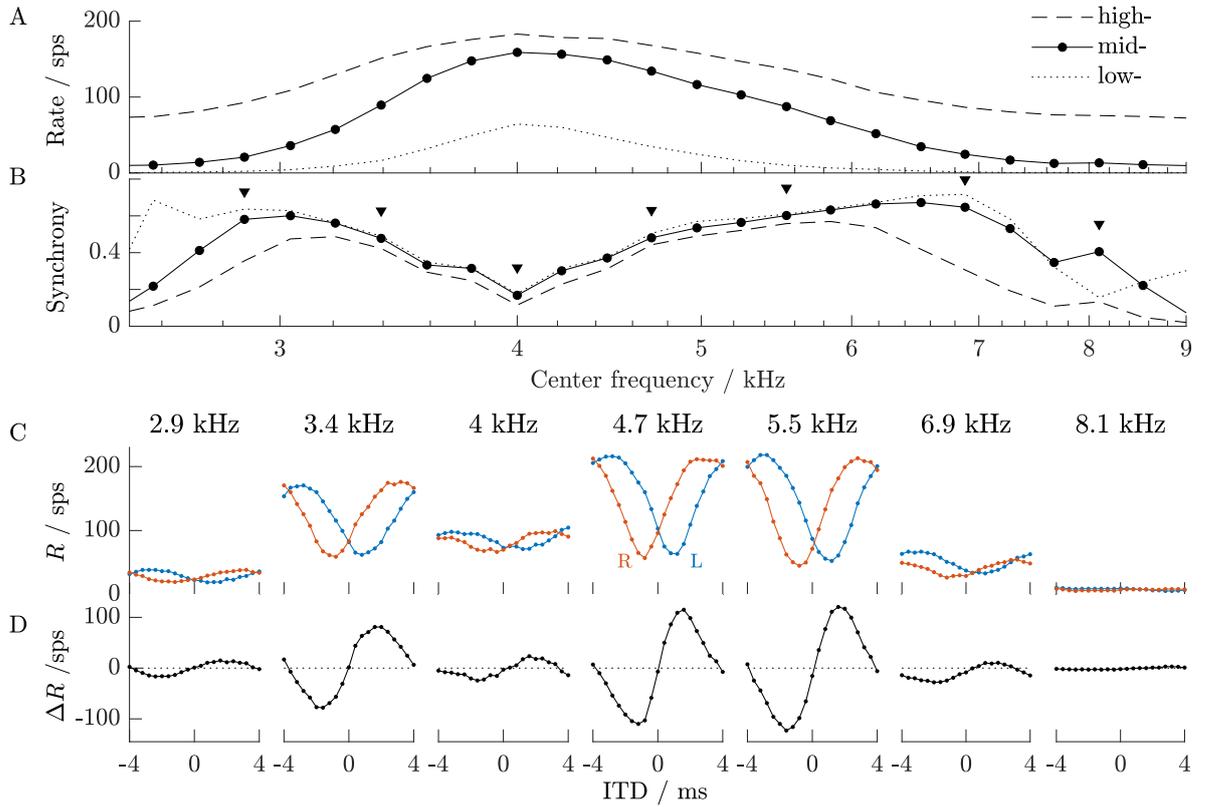
Figure 3.4.: Model responses across different CFs for a 68 dB SPL SAM tone with $f_m = 128$ Hz. (A) AN response rate. Different line styles indicate fiber types with different spontaneous rates. (B) Corresponding vector strength. Triangles indicate the center frequencies used in the panels below. (C) Rate-ITD functions for left (L) and right (R) hemisphere. (D) Rate difference (ΔR).

localization in cochlear implant users. With the focus of the current study on binaural interaction, we adopt this simple linear mapping of the mean hemispheric response rate difference to the extent of laterality, instead of employing a more complex mapping stage.

In previous psychoacoustic experiments, extents of laterality for high-frequency stimuli were commonly measured with an acoustic-pointing task (Bernstein and Trahiotis, 2003, 2012; Dietz *et al.*, 2015). In these experiments, listeners were first presented with the high-frequency target stimulus. They were then presented with a pointer stimulus, which was a band-limited Gaussian noise centered at 500 Hz having a bandwidth of 200 Hz. The listeners were asked to adjust the ILD of the pointer to match the perceived intracranial position of the target and that of the pointer stimulus. Pointer and target stimuli were repeatedly alternated until the subjects indicated that they had matched the position of the pointer and the target (open loop). The pointer ILD was then used as a measure for the extent of laterality. Positive and negative pointer ILDs indicate a right or left intracranial position, respectively.

Having thus obtained one neural response rate difference for each condition ($\Delta \overline{R}_i$), the last model stage relates the simulated $\Delta \overline{R}_i$ to the experimentally obtained pointer ILD. Assuming the simplest case of a linear relationship, a single scaling factor ρ connects the two quantities:

$$\hat{y}_i = \rho \times \Delta \overline{R}_i, \tag{3.4}$$

33

with $\hat{y}_i$ being the predicted pointer ILD for the ith condition. ρ is considered as a subject specific factor. To quantify the goodness of the prediction, the amount of variance accounted for (VAF,ψ) by the model was calculated by

$$\psi = 1 - \left[\sum_i (y_i - \hat{y}_i)^2\right] / \left[\sum_i (y_i - \overline{y})^2\right], \tag{3.5}$$

with y_i and $\hat{y}_i$ being the observed and predicted acoustic pointer ILDs, respectively, and $\overline{y}$ the mean value of all observed conditions. Additionally, we state the root-mean-square error ϵ.

3.3.5. Simulation design

The same stimuli as in the respective studies were generated. They differ from the original stimuli only in sampling rate (set to 100 kHz needed by the peripheral stage), duration, and ITDs. The ITDs were inserted after the peripheral processing to reduce computational demand. From the total 2.2 s stimulus duration, the first 200 ms were discarded to avoid stimulus onset effects and associated adaptation (of auditory nerve fibers). In the psychoacoustic experiment, an open loop for matching the pointer ILD was used. Therefore, it is unclear how many seconds the subjects exploited to adjust the pointer. Nevertheless, for such a pointer paradigm, the simulated duration is expected to be uncritical, influencing only the spike standard error but not the mean rates (this holds for the psychoacoustic responses as well).

In total, $M_{\mathrm{ex}} + M_{\mathrm{inh}}$ AN fibers were simulated for each of the $N = 30$ CFs in the range of 2 to 10 kHz distributed equidistantly along the tonotopical axis (according to Greenwood, 1961). This results in 840 AN fibers per hemisphere for default parameters merging into the decoding stage described above. For randomly selected AN fibers, each with a unique response rate, a left or right bias in overall activity was occasionally observed in the EI-stage, leading to a lateralization bias not observed in NH listeners. In a real system more neurons or more central stages are expected to average out or compensate any bias. To avoid such bias in our model, a deterministic distribution of spontaneous rates was generated, representative for the used fiber-type: We picked M_{ex} or M_{inh} r_{s}-values, equidistant on the cumulative Gaussian distribution of the according fiber-type. We assume $\mu = 4\,\mathrm{sp/s}$, $\sigma = 4\,\mathrm{sp/s}$ in a range of 0.5-18 sp/s for medium-spontaneous rate fibers and $\mu = 70\,\mathrm{sp/s}$, $\sigma = 30\,\mathrm{sp/s}$ in a range of 18-180 sp/s for high-spontaneous rate fibers (Liberman, 1978; Bruce *et al.*, 2018). The refractory times t_{rel} (131-894 μs) and t_{abs} (209-692 μs) were randomly picked for each AN-fiber, but identically in both hemispheres (for values and range see Miller *et al.*, 2001; Bruce *et al.*, 2018). The model was implemented in MATLAB (MathWorks, Natick, MA). The code is published as supplemental material to this paper.

3.4. Predicting lateralization of high-freqeuncy sounds

To fit the model output to the three experimental data sets, the focus was on varying the parameters of the EI-model stage. This stage has seven parameters: the numbers of excitatory and inhibitory inputs (M_{ex} and M_{inh}), duration of the excitatory window (W_{ex}), duration of the inhibitory window (W_{inh}), the response threshold (θ), an inhibitory gain factor (δ) to increase the weight of inhibitory inputs, and the duration of the refractory period (T) (Table 3.1).

Table 3.1.: Parameters of the EI-model (first column), respective symbols used (second column), the simulated range (third column), and parameters leading to the best performance obtained by the grid search (fourth column).

Parameter	Symbol	Simulated range	Best performance
Number of excitatory inputs	M_{ex}	18-22	20
Number of inhibitory inputs	M_{inh}	6-9	8
Excitatory window	W_{ex}	0.5-1.6 ms	1.1 ms
Inhibitory window	W_{inh}	2.5-4.0 ms	3.1 ms
Response threshold	θ	2-5	3
Inhibitory gain	δ	1-3	2
Refractory period	T	$T \leq W_{\text{ex}}$ [a]	1.6 ms

[a]The refractory period is set to 1.6 ms (or equal to W_{ex} if $W_{\text{ex}} > 1.6$ ms to avoid multiple output spikes resulting from the same input spike).

For a detailed description of each parameter and its effects, see Ashida *et al.* (2016). The refractory period was kept constant at 1.6 ms. In addition to varying these EI-stage parameters, the model was tested with either high- or mid-spontaneous rate AN fibers.

Furthermore, there is the subject specific factor ρ to relate the rate difference to pointer ILD. The aim of this study is to model only average data, but even the average ρ of small cohorts may vary. Therefore, ρ is fitted to the mean data of each of the three data sets respectively.

The aim of the parameter variation was to account for most of the variance in three published data sets. Figure 3.5 gives an overview of how the model performance depends on various parameters for each of the three studies presented below. There is no common optimum set across studies, but many maxima occur within the parameter test range that was inspired by physiology. A range of parameters allows to account for much of the variance in all studies. The parameter set listed in Table 3.1 with $W_{\text{inh}} = 3.1$ ms (cf. filled symbols in Fig. 3.5) explains much of the variance of all three data sets and therefore was chosen to generate the simulation of data from the three studies shown in all figures below.
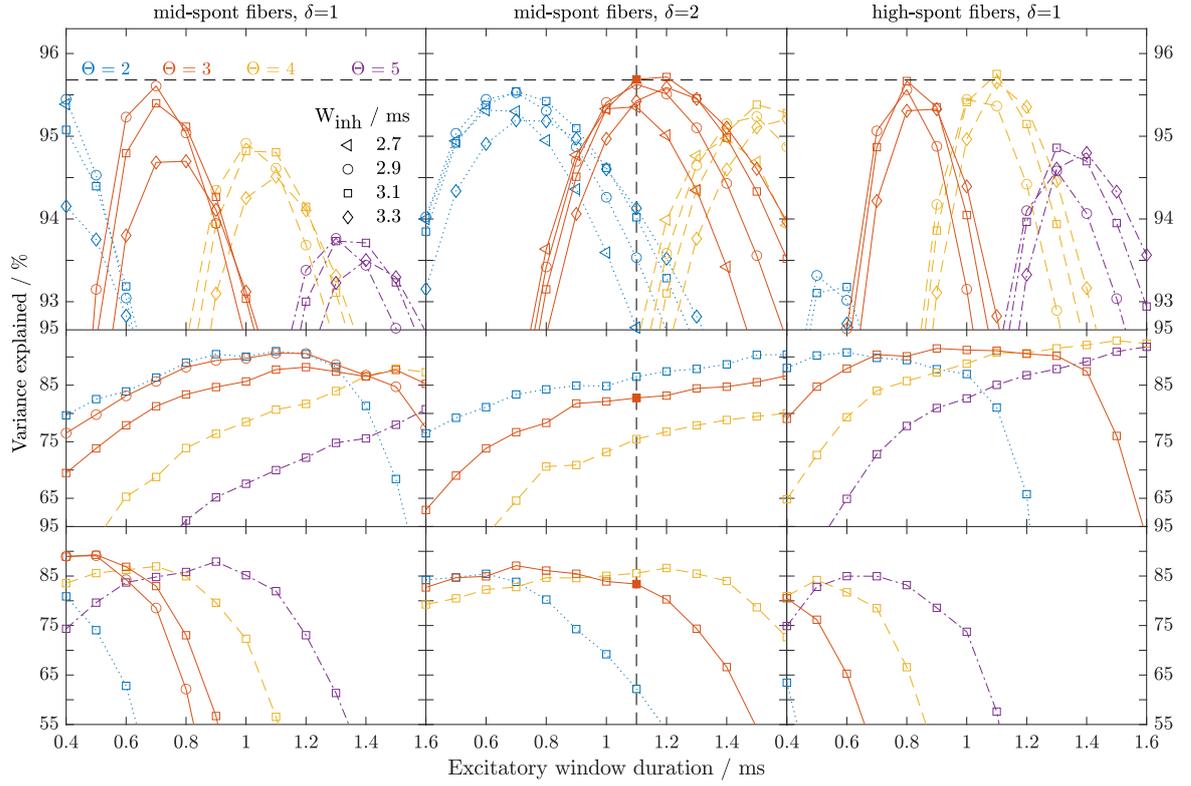
Figure 3.5.: Variance explained (ψ) for Bernstein and Trahiotis (2012) with ρ=0.29 dB/sps (top row), Bernstein and Trahiotis (2003) with ρ=0.48 dB/sps (middle) and Dietz et al. (2015) with ρ=1.02 dB/sps (bottom row). Color and line style indicate the threshold: $\theta = 2$ (blue, dotted), $\theta = 3$ (red, solid), $\theta = 4$ (yellow, dash-dotted), $\theta = 5$ (purple, dashed). Note the different ordinate ranges. First column is with mid-spontaneous rate fiber input and an inhibitory gain $\delta = 1$, second column with $\delta = 2$ and the third column represents model predictions with high-spontaneous rate fibers projecting to the EI stage and $\delta = 1$. The filled symbols represent for the chosen parameter combination (table 3.1) for the particular data sets.

3.4.1. Raised-sine stimuli

To investigate the influence of envelope shape on the extent of laterality, previous studies (e.g., Bernstein and Trahiotis, 2012) used raised-sine stimuli. These stimuli are based on SAM tones and allow for an arbitrary exponent (n) that influences the peakedness of the stimulus envelope independent of f_m and modulation depth (m_d) [John *et al.* (2002); see Fig. 1 of Bernstein and Trahiotis (2009) for examples of raised-sine stimuli]. An exponent of $n = 1$ generates a conventional SAM tone. For higher exponents, the raised sine has a steeper slope. In a compact form (Bernstein and Trahiotis, 2012), the stimulus is defined by

$$y(t) = a \, \sin(2\pi f_c t) \left[(1 - m_d) + 2m_d \, (\sin(\pi f_m t))^{2n} \right]. \tag{3.6}$$

In the study of Bernstein and Trahiotis (2012), f_c was fixed at $4\,\text{kHz}$, while modulation frequencies of 32, 64, 128 and 256 Hz, and modulation depths of 0.25, 0.5, 0.75 and 1.0 were applied. The stimulus was either raised to a power of $n = 1$, or $n = 8$. ITDs of 0, 200, 400, 600, 800, and $1000\,\mu\text{s}$ were employed and ILDs of -8, -4, 0, 4, and $8\,\text{dB}$ were applied by symmetrically varying the sound pressure level (in dB) on the left and the right side. An overall sound level of $68\,\text{dB}\,\text{SPL}$ was used. All possible stimulus parameter combinations were tested, resulting in a total of 960 stimulus conditions (Fig. 3.6). This extensive data set captures many factors influencing sound lateralization, and is used here as the central dataset for fitting the model parameters of the EI-stage. Arguably because of its richness, it restricts the acceptable range of parameters more than the other data sets shown in Fig. 3.5.

Model predictions for all stimulus conditions are displayed in Fig. 3.6 as solid lines (after converting rate difference to pointer ILD). The different stimulus conditions are organized into 32 panels for all the combinations of the parameters f_m, m_d and n of the raised sine. Each panel shows the acoustic pointer depending on the ITD and ILD. The mean measured values of acoustic pointer (Bernstein and Trahiotis, 2012) are shown as symbols. The steepness of the relation between ITD and pointer ILD in each panel can be used as an indicator for pronounced ITD-based lateralization.

Figure 3.6.: Extents of laterality experienced by listeners (Bernstein and Trahiotis, 2012) are shown as black symbols. They are grouped based on the parameters of the stimulus. The predictions of the model (scaling factor $\rho = 0.29\,\text{dB/sps}$) are shown as solid lines with the ITD as abscissa. The four blocks (each consisting of eight panels) represent the four modulation frequencies, the panels in columns organize the different depths of modulation, while the two different exponents are separated into rows. Each panel shows the acoustic pointer in dB on the ordinate, which quantifies the extent of laterality as a function of ITD. Five different ILDs are used in each panel. The variance accounted for is stated in each panel separately. Adapted and redrawn from Fig. 1 of Bernstein and Trahiotis (2012) with permission.
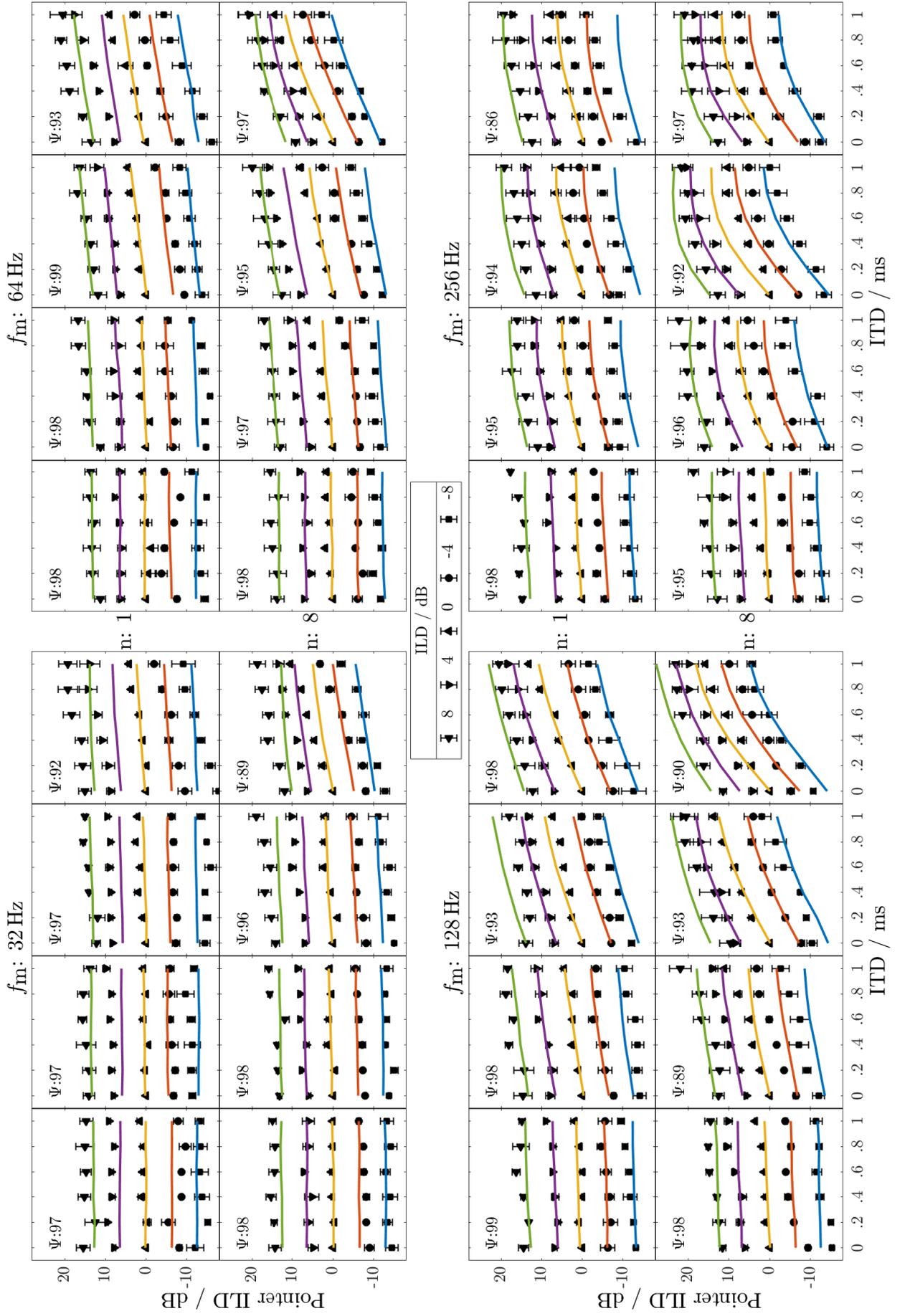
3. Neural rate difference model can account for lateralization of high-frequency stimuli

Taken together, three main effects are apparent in both the observed data and the model predictions (Fig. 3.6): (1) The ILD-based lateralization is almost constant; i.e, the five lines in each panel primarily differ by an offset which is almost constant across panels, with one exception described below. (2) The extent of ITD-based lateralization increases with modulation depth and with modulation frequency up to 128 Hz but then decreases slightly for modulation frequencies of $f_m = 256$ Hz. (3) The relation between ITD and pointer ILD is linear for small modulation frequencies, combined with small depth of modulation ($f_m \leq 64$ Hz and $m_d \leq 0.5$), but the relation becomes curved for high modulation frequencies and full modulation ($f_m \geq 128$ Hz and $m_d = 1.0$).

The model was able to quantitatively reproduce all three main effects and accounted for 95.7 % of the variance in the data ($\epsilon = 2$ dB) with the parameter set chosen here (see Table 3.1). A model test using only 1-s signals, rather than 2 s, accounted for an almost identical 95.5 % of the variance. Because the chosen metric captures only the overall trends in the data, we also state the explained variance separately for each panel. Two deviations are apparent. First, for an exponent of $n = 8$, $f_m = 32$ Hz, and full modulation (Fig. 3.6, upper left block, panel in the lower right corner) the model underestimates lateralization when both ITD and ILD are favoring the same direction. Secondly, for $f_m = 256$ Hz, $n = 1$, and full modulation, the model underestimates the impact of a positive ITD when a negative ILD is present at the same time (Fig. 3.6, lower right block, panel in the upper right corner).

The model output for all 960 conditions (prior to across-frequency integration and conversion into acoustic pointer) is shown in Fig. 3.7. The rate differences at each CF are related to the lateralization data recorded by Bernstein and Trahiotis (2012). The rate difference responses of off-frequency neurons show a higher correlation with the observed data. For neurons with a CF matching the carrier frequency of 4 kHz, the model accounted for only 30 % of the variance in the data, primarily ILD-based lateralization. In contrast, the off-frequency pair of model neurons accounts for 92 % of the variance.
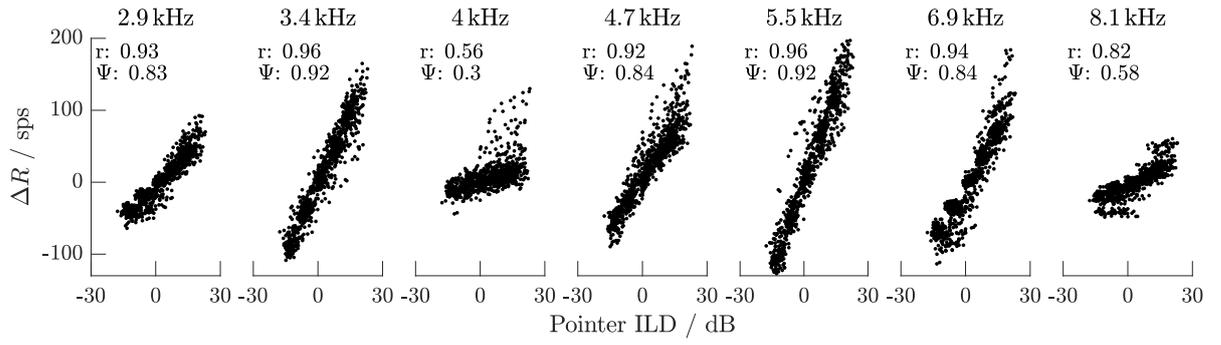
Figure 3.7.: Scatter plots of the rate difference (ΔR) between left and right EI-model neuron pairs versus observed pointer ILDs (Bernstein and Trahiotis, 2012) for all conditions and across seven different example CFs. In each plot, 960 single data points are shown, corresponding to the varied parameter combinations of the high-frequency raised sines. The correlation coefficient (r) and the variance accounted for (ψ) are given for each of the seven positions in the tonotopic array. The model was parameterized with the best overall parameter set (see Table 3.1).

3.4.2. Transposed stimuli

Transposed stimuli were designed to better mimic low-frequency tonal responses in high-frequency AN fibers. Low-frequency tones depolarize the hair cells only during the condensation phase, i.e., about 50 % of each period. In contrast, sinusoidally amplitude modulated high-frequency stimuli cause a continuous depolarization of hair cells, except for the short moments of zero amplitude where the receptor potential draws near the resting potential.

The purpose of transposed stimuli is to mimic the activation pattern of a low-frequency stimulus in a high-frequency region. To do so, a low-frequency base stimulus (e.g., a low frequency sinusoidal) is half-wave rectified and subsequently low-pass filtered, to simulate the functional role of hair cells. The output serves as the modulator and is multiplied by a high-frequency carrier (van de Par and Kohlrausch, 1997). Once processed by the real hair cells, the output provides high-frequency AN fibers with a temporal excitation pattern that is relatively similar to that generated by the base stimulus available to low-frequency AN fibers. A difference remains in the rarefaction phase where the base stimulus causes hyperpolarization but the transposed stimulus results in a resting potential.

Colburn and Esquissaud (1976) hypothesized that the similar AN activation should cause similar binaural interaction, i.e. for low-frequency and transposed tones. In contrast, conventional high-frequency stimuli are expected to produce weaker interaural differences. The hypothesis was tested by Bernstein and Trahiotis (2003) comparing three types of stimuli: low-frequency noise, transposed noise, and high-frequency narrow-band Gaussian noise. To test our high-frequency model, we used the latter two stimulus types (Fig. 3.8 B). The stimuli were generated in the same way as in Bernstein and Trahiotis (2003) and the carrier frequency for all stimuli was fixed at 4 kHz, while bandwidths of 25, 50, 100, 200, and 400 Hz where used. The transposed stimulus was modulated with a half-wave rectified, low-frequency noise centered at either 125 Hz or 250 Hz. For the 125 Hz center frequency stimulus, the largest bandwidth was 200 Hz.

Figure 3.8 B shows the stimuli used for our simulations: (1) transposed low-frequency narrow-band Gaussian noise centered at 125 Hz (Fig. 3.8 B, upper panel), (2) transposed low-frequency narrow-band Gaussian noise centered at 250 Hz (Fig. 3.8 B, middle panel), (3) high-frequency narrow-band Gaussian noise (Fig. 3.8 B, bottom panel). The stimulus level was set to 72 dB SPL as in Bernstein and Trahiotis (2003).

The model predictions were compared to the psychoacoustic data taken from the corresponding study (Bernstein and Trahiotis, 2003). EI-model parameters were kept unchanged (Table 3.1) for the data shown in Fig. 3.8. Only a different scaling factor was fitted [Eq. (3.4)].

In general, larger extents of laterality were observed with increasing ITD (Fig. 3.8 A). The bandwidth was the dominant factor for the lateralization of high-frequency Gaussian noise. Three main effects could be observed from the psychophysical data (Fig 3.8 A, symbols): (1) For high-frequency Gaussian noise, there was virtually no lateralization for bandwidths of 25 Hz and 50 Hz (Fig. 3.8 A, top two panels), and the lateralization increased with increasing bandwidth (Fig. 3.8 A, solid lines). (2) The transposed stimuli produced larger extents of laterality compared to non-transposed stimuli (Fig. 3.8 A, high-frequency Gaussian noise compared to transposed 125 Hz and 250 Hz stimuli) as the acoustic pointer was adjusted by listeners with higher values. (3) The pointer ILDs were similar for the two different center modulation frequencies of the transposed noises (Fig. 3.8 A, transposed 125 Hz and 250 Hz stimuli). Our model reproduced most of these trends in the data (Fig. 3.8 A) and accounted for 82.7 % of

the variance (ϵ=2.4 dB). Other parameter sets, more optimal for this particular study allow to account for just over 90 % of the variance (see Fig. 3.5).
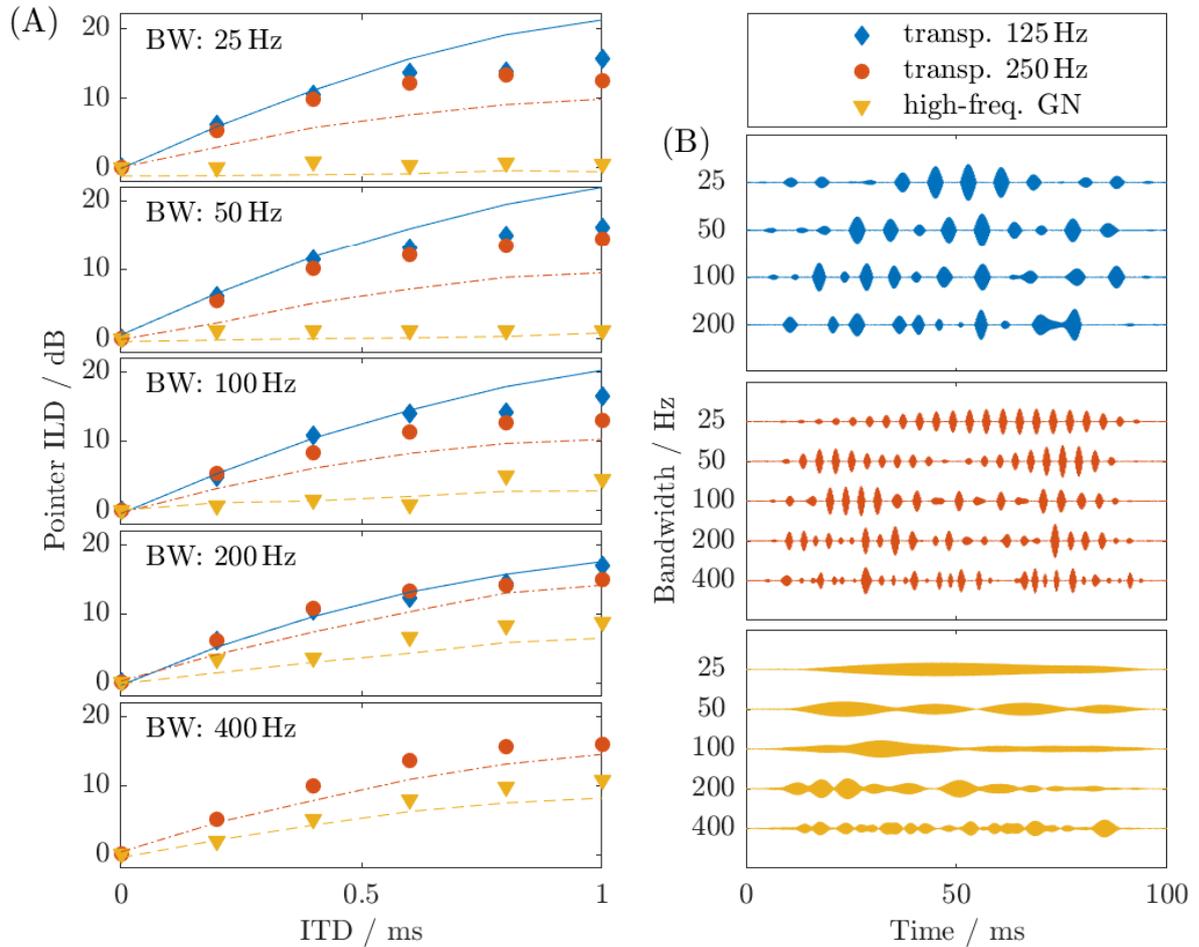
Figure 3.8.: Stimulus conditions, psychoacoustic data and model predictions. **(A)** Psychoacoustic data measured by Bernstein and Trahiotis (2003) represented as symbols, and our model predictions shown as curves (scaling factor $\rho = 0.48$ dB/sps). Data adapted and redrawn from Fig. 2 of Bernstein and Trahiotis (2003) with permission. The input was transposed noise centered at 125 Hz (densely dash-dot-dotted, diamond), transposed noise centered at 250 Hz (solid, bullet), and high-frequency Gaussian noise (dash-dotted, down-facing triangle). **(B)** High-frequency transposed tones as counterparts of low-frequency Gaussian noise with different center frequencies of either 125 Hz (top) or 250 Hz (middle) and different bandwidths (BW), and high-frequency Gaussian noise with a center frequency of 4 kHz (bottom).

3.4.3. Envelope rise- and decay elements

Another class of high-frequency AM stimuli was constructed by independently varying durations of the rising envelope segment (rise time), the falling envelope segment (decay), the pause between lobes, and of the peak plateau of the temporal envelope (e.g., Klein-Hennig *et al.*, 2011; Dietz *et al.*, 2015).

In particular, this configuration allows generation of temporally asymmetric envelopes. The extent of ITD-based lateralization perceived by subjects was found to be particularly high, when a short rise time and a non-zero pause were combined (Dietz *et al.*, 2015). The steepness of the rising part of the envelope had more influence on lateralization than the steepness of the decay. So far, such differences have not been accounted for with cross-correlation based models (e.g., Klein-Hennig *et al.*, 2011).
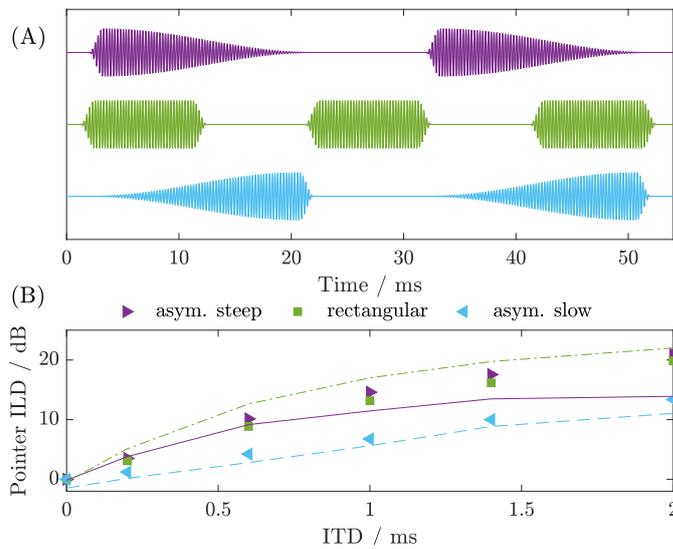
Figure 3.9.: Stimulus conditions, psychoacoustic data and model predictions. **(A)** High-frequency stimuli with different envelope shapes (Dietz *et al.*, 2015). Envelope shapes have a short 1.25 ms rise time, 18.75 ms decay and 0 ms plateau (top, right-facing triangle), a short 1.25 ms rise time, 8.75 ms plateau and 1.25 ms decay (middle, square), and a long 18.75 ms rise time, 0 ms plateau and 1.25 ms decay (bottom, left-facing triangle). **(B)** The corresponding psychoacoustic data (Dietz *et al.*, 2015) represented as symbols. Line styles for model predictions with a continuous axis of ITD were dotted for a short rise time and shallow decay, dashed for rectangular modulation, and densely dash-dotted for a long rise time and fast decay. Adapted and redrawn from Fig. 2 of Dietz *et al.* (2015) with permission. For reasons of comparability the pointers were shown not normalized with respect to the subject's average pointer ILD (scaling factor ρ =1.02 dB/sps).

The current model was tested with three different envelope shapes: (1) a short rise time with zero plateau and long decay (Fig. 3.9A, top), (2) a short rise time with a pause-equivalent plateau duration and a fast decay (Fig. 3.9A, middle), and (3) a long rise time with zero plateau and fast decay (Fig. 3.9A, bottom). The carrier was a fully modulated 4 kHz tone, matching the peak level of a 65 dB SAM as in Dietz *et al.* (2015). Again, the model with parameters of best performance (Table 3.1) was used and a new scaling factor ρ was calculated [Eq. (3.4)] to account for the different subject group and differences in the experimental setting. Primarily due to the lower stimulus level in this study the scaling factor had to be larger compared to the

studies of Bernstein and Trahiotis (2003, 2012).

The model predictions were compared with the data from the normal-hearing listeners in Dietz *et al.* (2015). In both data and model predictions, a long rise time and short decay led to the least pronounced lateralization (Fig. 3.9 B, left pointing triangles and dash-dotted line) and to a linear increase of lateralization with ITD up to 2 ms. Also, the model correctly predicted the 50-80 % higher lateralization for the condition with short rise and short decay. However, the model clearly underestimated the lateralization for a short rise time and long decay (Fig. 3.9 B, squares and dotted line), particularly at large ITDs. The model accounted for 83.4 % of the variance (ϵ=2.8 dB).

3.5. Discussion

The present study aimed to develop a rate difference model that links the physiology of mammals with the observed binaural phenomena in human psychoacoustics. We proposed a computational model that captures the effect of ILD- and envelope-ITD-based sound lateralization of narrowband high-frequency stimuli. At a given sound level, predicted lateralization is proportional to the summed hemispheric rate difference of EI- model neurons with identical parameters but different CFs. The EI-model neurons resemble the binaural processing core, corresponding to the LSO (Ashida *et al.*, 2016), which is the primary nucleus in the mammalian brain for encoding both ILD and envelope ITD (Tollin, 2003).

3.5.1. Influence of model parameters

The EI-model (Ashida *et al.*, 2016) has seven parameters that were never fitted to account for human perception. Such a number of free parameters may raise concerns about potentially being able to fit to any data set with up to seven independent stimulus dimensions. Nevertheless, large covariances of parameter influences can be expected to limit the degrees of freedom and the dataset from Bernstein and Trahiotis (2012) which comprises a 5-dimensional stimulus space is well suited to study the model parameter dependences.

The data from Fig. 3.5 (top row) reveal that for most ad-hoc chosen sets of parameters only the duration of the excitatory window W_{ex} had to be optimized to account for at least 94 % of the variance, just as the original model from Bernstein and Trahiotis (2012). Similarly, for any fixed W_{ex} in the tested range (see Table 3.1), parameters such as gain, threshold, or fiber number could be adjusted to obtain an excellent fit. This observation suggests strong interdependences across the parameters. For example, decreasing the number of input fibers has a similar effect as increasing the threshold. Similarly, halving the inhibitory gain is comparable to doubling the number of inhibitory input fibers.

Only the duration of the inhibitory input had to be within a narrow range between 2.9 and 3.3 ms, independent of the values of θ, W_{ex} and M_{inh}. With $M_{\mathrm{inh}} = 3.1$ ms fixed, at least 94.8 % of the variance can be explained by optimizing only the duration of the excitatory input, irrespective of other parameters (within the tested range). M_{inh} has a critical role in determining the upper modulation frequency limit for ITD sensitivity and the starting decline in ITD-based lateralization around 256 Hz. For subjects with a higher frequency limit (e.g., Monaghan *et al.*, 2015), a shorter integration window would be required.

If the focus is put on sparse transient stimuli, i.e., low f_{m}, $m_{\mathrm{d}} = 1$, and a long modulation

trough, as in Dietz *et al.* (2015), the influences of the parameters are somewhat different (Fig. 3.5). The short modulation onset events sometimes cannot cause similarly large response rates as for other stimuli with the same lateralization. Overall responses are weaker and the simple rate difference metric may underestimate the influence of the few reliable responses. The same effect has been observed for the $f_\mathrm{m} = 32\,\mathrm{Hz}$, $m_\mathrm{d} = 1$, $n = 8$ conditions of Bernstein and Trahiotis (2012), where our model underestimates ITD-based lateralization. In summary, it appears as if the model utilizes two central degrees of freedom: 1) Inhibitory integration time and 2) the right combination of excitation, inhibition, and threshold.

3.5.2. Relation to other models

The main difference to other models is arguably in the structure: While other models have fully or partially separated mechanisms to encode ILDs and envelope ITDs (e.g., Bernstein and Trahiotis, 2012; Takanen *et al.*, 2014), the EI-model neurons can simultaneously account for both ILD- and envelope ITD-based extents of laterality. This is a more constrained situation for a model and at the same time in closer relation to neural responses in the LSO (Tollin, 2003).

With only two free EI-model parameters (duration of excitatory and inhibitory window), the model can account for about 95 % of the variance across a wide range of values for the other parameters in the biggest data set. For reference, the delay line-based model of Bernstein and Trahiotis (2012) accounted for a practically identical 94 % of the variance in the same data set. The overall high variance accounted for by all models, is to some extent due to the fact that already a pure ILD-based lateralization model accounts for 63.1 % of the variance in this particular data set. The overall performance of the two models is similar, but some differences can be observed. For instance, at $f_\mathrm{m} = 256\,\mathrm{Hz}$, $n = 8$, and $m_\mathrm{d} = 0.5$, the delay-line-based model predicted a linear dependence of lateralization and envelope ITD and accounted for 75.0 % of the variance in the data, while our model accounted for 96.4 % in the respective panel (Fig. 3.6). In contrast, the proposed model underestimated ITD-based extents of laterality at $f_\mathrm{m} = 32\,\mathrm{Hz}$, $n = 8$, and full modulation ($\psi = 89\,\%$), while the model of Bernstein and Trahiotis (2012) accounted for 97.8 % of the variance in these conditions. Lastly, the proposed model even captured the curved relation at high modulation frequencies with high exponent (Fig. 3.6, bottom right block, lower panels).

Employing identical EI-parameters with a different scaling factor ρ to high-frequency transposed and high-frequency Gaussian noise of different bandwidths, our model accounted for 82.7 % of the variance in the psychoacoustic data of Bernstein and Trahiotis (2003). The model correctly captured the dominant effects in this data set, while it underestimated the lateralization of non-transposed noise at large bandwidths. The original cross-correlation model used to simulate this data (Bernstein and Trahiotis, 2003) overestimated lateralization for high-frequency Gaussian noise by up to 10 dB pointer ILD at all bandwidths, except at 25 Hz. Our model correctly predicts the much smaller lateralization for Gaussian noise, because only the transposed-noise condition facilitates high synchronous excitation to steep modulation onsets, after inhibition has ceased during short intervals of zero amplitude. The same feature causes the more pronounced lateralization in the EI-model for envelopes with steep onsets compared to shallow onsets and steep offset in section 3.4.3.

3.5.3. Off-frequency integration

In experimental data (e.g., Joris and Yin, 1992), as well as in simulated spike trains of the peripheral model (Bruce *et al.*, 2018), the spike rate and the phase locking of AN fibers depend greatly on the stimulus level. Physiologically, AN fibers show the best phase locking to envelopes at a sound pressure level of about 20 dB, and the degree of phase locking decreases at higher levels (Joris and Yin, 1992). This response characteristic is also captured by the peripheral stage of the model (see Fig. 13 in Zilany *et al.*, 2009). While intermediate levels of phase locking are still sufficient to facilitate ITD-based lateralization, the current model does not account for both ITD and ILD based lateralization in a quantitative manner at CF = f_c. In contrast, envelope ITD-based lateralization is level independent in the range from 45 to 65 dB SPL (Dietz *et al.*, 2015), and detection sensitivity even improves with increasing level (e.g., Bernstein and Trahiotis, 2008). We accounted for this discrepancy by employing a population of spiking auditory model neurons with a broad range of CFs. This is an implementation of off-frequency hearing, as suggested by Dreyer and Delgutte (2006). Incorporating these off-frequency components crucially enabled the model to explain the data with high accuracy (Figs. 3.4 and 3.7). However, on-frequency neurons were also necessary to account for the most variance, because they were instrumental for stimulus conditions with low modulation depths.

To investigate the role of off-frequency hearing, notched-noise stimuli are commonly used to mask information in off-frequency channels. Bernstein and Trahiotis (2008) report that notched noise has only a modest influence on threshold envelope ITDs. In the periphery model (Bruce *et al.*, 2018) notched noise appears to improve on-frequency phase locking, suggesting off-frequency hearing in the absence and on-frequency hearing in the presence of notched noise. Without published extent of laterality data with notched noise, a quantitative analysis is beyond the scope of the present study. It is expected, however, that the model requires a more sophisticated back-end to operate in the presence of any interfering noise.

Peripheral band-pass filtering may further contribute to the importance of off-frequency channels on envelope ITD perception. The amplitude modulation after filtering can be more pronounced for off- compared to on-frequency channels, especially for high modulation frequencies (Monaghan *et al.*, 2015).

3.5.4. Decoding EI output and onset dominance

A variety of different decoding stages has been suggested in previous studies (for review see Dietz *et al.*, 2018). One common simplistic approach is to decode extent of laterality through a weighted average across frequency channels (Bernstein and Trahiotis, 2012; Takanen *et al.*, 2014; Kelvasa and Dietz, 2015). A linear mapping as proposed for simple stimuli by Kelvasa and Dietz (2015) was employed in our model as a particularly simple decoder option. In its current implementation the model simply averages over the 30 CFs. However, the broadness of neural excitation depends critically on overall level. Therefore, the current decoding stage is not expected to operate in the presence of any interferer or to produce level independent lateralization as reported at moderate signal levels (Dietz *et al.*, 2015). The scaling factor ρ can be varied to compensate for the vast changes of response rate differences with overall level. A rate ratio could be an alternative to a rate difference metric, that appears to be less level dependent but underestimates the influence of differences at large EI rates. Some transition

between the two metrics or a decoding stage mapping the two rates in a more complex way to the extent of laterality appears to be a promising next step. Another possibility is to normalize the hemispheric rate difference based on AN response rates (Encke and Hemmert, 2018).

The complexity of the *decoding* stage should be linked to the complexity of subcortical ITD *encoding*. Some previous studies reported a diversity of neural responses to envelope ITDs along the auditory pathway. For example, recordings from the inferior colliculus (IC) of guinea pigs revealed step-type ITD sensitivity alongside steep trough-type ITD-rate functions and more gradual functions - all in response to the same envelope shape (Dietz *et al.*, 2016; Greenberg *et al.*, 2017). A more diverse population of binaural interaction neurons may not only lead to an improved model performance, it may also give a more realistic representation of the real-world biological system. Specifically, Dietz *et al.* (2016) required onset-type input to model binaural model neurons. The onset-type input was facilitated by a simplistic CN stage and was necessary to model the contrast between very pronounced ITD tuning with short sharp onsets compared to a lack of ITD tuning with long gradual onsets, observed in some IC neurons. Onset-type input was not necessary in the present study to account for the data of Bernstein and Trahiotis (2003, 2012). In simulating the different extent of laterality generated by the short and long onsets, the present simple model performs better than cross-correlation-based models (Klein-Hennig *et al.*, 2011) but still clearly underestimate the difference (Fig. 3.9 B). A more diverse mix of model neurons with some modulation onset-dominated specimen is expected to be useful to quantitatively account for the data. Thus, a diverse pool combination of differently behaving model neurons (e.g., Dietz *et al.*, 2016) appears to be supported by both physiology and perception.

Introducing various types of neurons, however, would make the parameter optimization very difficult and require a more complex decoding stage. A complex decoding stage could further estimate the relative importance of the different channels across the tonotopic array. It could also use both left and right rates rather than just the rate difference to extract more information and to be more robust against other stimulus variations, such as overall level. Nevertheless, even the linear decoder can directly relate the output of binaural model neurons to extent of laterality at a fixed sound level. This is in contrast to the decoding stages of several other models that estimate the input ITDs or similar physical quantities, rather than perceptual measures (e.g., Goodman *et al.*, 2013; Encke and Hemmert, 2018). Such a calculation of input ITD will not be useful for predicting the lateralization of complex sounds from the present study, because the same envelope ITD leads to substantially different extents of laterality (Figs. 3.6, 3.8 and 3.9).

3.5.5. Physiological, pathophysiological, and anatomical considerations

Although our model is arguably no less complex than other models of envelope-ITD perception (e.g., Cai *et al.*, 1998), from a physiological standpoint the structure is still highly simplified. The simulated model neurons of the AN stage provide direct input to the stage of binaural interaction, bypassing all other brainstem structures such as the cochlear nucleus. Nonetheless, the model produces ITD or ILD rate functions that match functions obtained experimentally in the LSO, with model parameters that are within a physiologically realistic range (compare to, e.g., Sanes, 1990; Tsuchitani, 1988).

The model is further simplistic in only considering AN fibers with either medium- or high-

spontaneous rates. Combining different fiber types or including a cochlear nucleus stage would increase the number of parameters and add complexity. Figure 3.5 revealed that the model is generally robust against the choice of fiber type. With high-spontaneous rate fibers the model accounts for an identical 95.7 % of the variance in the biggest data set.

Overall, the model operates between the typical forces of physiological accuracy versus being manageable, interpretable, and therefore rather simplistic (Wilson and Collins, 2019). The ultimate design choice was inspired by Colburn (1973). He listed "attractive characteristics" for models of retrocochlear processing that form the basis of many successful auditory models: (1) the inputs are auditory-nerve responses (2) the processing is not unreasonable for neural structures and (3) quantitative predictions can be derived. Thus, we connected an established front-end AN model with a relatively simplistic but physiologically plausible EI-stage and the most simplistic rate-difference decoding back-end.

The above-mentioned co-dependence of model parameters and ability to construct well-performing models with only two independent fitting parameters at the stage of binaural interaction is also interesting from a pathophysiologic or audiological perspective. The abundance of highly co-varying parameters allows a system to compensate for any suboptimal spatial mapping caused mild periphery impairments. It may be that, e.g., an AN fiber loss is compensated for by a reduction of inhibitory gain (e.g., Schaette and McAlpine, 2011) to keep a very similar lateralization performance. However, this hypothesis must be interpreted with caution and would require further testing.

3.6. Summary and conclusion

The present study demonstrated that the lateralization of complex, high-frequency stimuli can be simulated with a relatively simplistic model deduced from mammalian auditory brainstem physiology and a linear hemispheric rate-difference decoder. To summarize:

1. Only one pair of model EI-neurons (composed of one neuron from the left and one from the right) was employed for each center frequency.

2. One neuron pair simultaneously encodes both ILD and envelope ITD, so that the rate difference is proportional to the extent of laterality at a given sound level.

3. Off-frequency model units are essential for envelope ITD-based lateralization at the sound levels commonly used in psychoacoustic experiments in the absence of notched noise.

4. The model accounts for 95.7 %, 82.7 %, and 83.4 % of the variance in three data sets using the same set of parameters. Lateralization of stimuli with 2 ms ITD can be simulated without delay lines.

3.7. References

Ashida, G., Kretzberg, J., and Tollin, D. J. (**2016**). "Roles for coincidence detection in coding amplitude-modulated sounds," PLOS Computational Biology **12**(6), 1–27, doi: `10.1371/journal.pcbi.1004997`.

Ashida, G., Tollin, D. J., and Kretzberg, J. (**2017**). "Physiological models of the lateral superior olive," PLOS Computational Biology **13**(12), 1–50, doi: `10.1371/journal.pcbi.1005903`.

Bernstein, L. R., and Trahiotis, C. (**2003**). "Enhancing interaural-delay-based extents of laterality at high frequencies by using 'transposed stimuli'," The Journal of the Acoustical Society of America **113**(6), 3335–3347, doi: `10.1121/1.1570431`.

Bernstein, L. R., and Trahiotis, C. (**2008**). "Discrimination of interaural temporal disparities conveyed by high-frequency sinusoidally amplitude-modulated tones and high-frequency transposed tones: Effects of spectrally flanking noises," The Journal of the Acoustical Society of America **124**(5), 3088–3094, doi: `10.1121/1.2980523`.

Bernstein, L. R., and Trahiotis, C. (**2009**). "How sensitivity to ongoing interaural temporal disparities is affected by manipulations of temporal features of the envelopes of high-frequency stimuli," The Journal of the Acoustical Society of America **125**(5), 3234–3242, doi: `10.1121/1.3101454`.

Bernstein, L. R., and Trahiotis, C. (**2012**). "Lateralization produced by interaural temporal and intensitive disparities of high-frequency, raised-sine stimuli: Data and modeling," The Journal of the Acoustical Society of America **131**(1), 409–415, doi: `10.1121/1.3662056`.

Bruce, I. C., Erfani, Y., and Zilany, M. S. (**2018**). "A phenomenological model of the synapse between the inner hair cell and auditory nerve: Implications of limited neurotransmitter release sites," Hearing Research **360**, 40–54, doi: `10.1016/j.heares.2017.12.016` computational models of the auditory system.

Cai, H., Carney, L. H., and Colburn, H. S. (**1998**). "A model for binaural response properties of inferior colliculus neurons. ii. a model with interaural time difference-sensitive excitatory and inhibitory inputs and an adaptation mechanism," The Journal of the Acoustical Society of America **103**(1), 494–506, doi: `10.1121/1.421130`.

Carr, C., and Konishi, M. (**1990**). "A circuit for detection of interaural time differences in the brain stem of the barn owl," Journal of Neuroscience **10**(10), 3227–3246, doi: `10.1523/JNEUROSCI.10-10-03227.1990`.

Colburn, H. S. (**1973**). "Theory of binaural interaction based on auditory-nerve data. i. general strategy and preliminary results on interaural discrimination," The Journal of the Acoustical Society of America **54**(6), 1458–1470, doi: `10.1121/1.1914445`.

Colburn, H. S. (**1977**). "Theory of binaural interaction based on auditory-nerve data. ii. detection of tones in noise," The Journal of the Acoustical Society of America **61**(2), 525–533, doi: `10.1121/1.381294`.

Colburn, H. S., and Esquissaud, P. (**1976**). "An auditroy-nerve model for interaural time discrimination of high-freqeuncy complex stimuli," The Journal of the Acoustical Society of America **59**(S1), S23–S23, doi: `10.1121/1.2002503`.

Dietz, M., Klein-Hennig, M., and Hohmann, V. (**2015**). "The influence of pause, attack, and decay duration of the ongoing envelope on sound lateralization," The Journal of the Acoustical Society of America **137**(2), EL137–EL143, doi: `10.1121/1.4905891`.

Dietz, M., Lestang, J.-H., Majdak, P., Stern, R. M., Marquardt, T., Ewert, S. D., Hartmann, W. M., and Goodman, D. F. (**2018**). "A framework for testing and comparing binaural models," Hearing Research **360**, 92–106, doi: `10.1016/j.heares.2017.11.010` computational models of the auditory system.

Dietz, M., Wang, L., Greenberg, D., and McAlpine, D. (**2016**). "Sensitivity to interaural time differences conveyed in the stimulus envelope: Estimating inputs of binaural neurons through the temporal analysis of spike trains," The Journal of the Association for Research in Otolaryngology **17**, 313–330, doi: `10.1007/s10162-016-0573-9`.

Dreyer, A., and Delgutte, B. (**2006**). "Phase locking of auditory-nerve fibers to the envelopes of high-freqeuncy sounds: Implications for sound localization," Journal of Neurophysiology **96**(5), 2327–2341, doi: `10.1152/jn.00326.2006`.

Encke, J., and Hemmert, W. (**2018**). "Extraction of inter-aural time differences using a spiking neuron network model of the medial superior olive," Frontiers in Neuroscience **12**, 140, doi: `10.3389/fnins.2018.00140`.

Goldberg, J. M., and Brown, P. B. (**1969**). "Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization.," Journal of Neurophysiology **32**(4), 613–636, doi: `10.1152/jn.1969.32.4.613`.

Goodman, D. F., Benichoux, V., and Brette, R. (**2013**). "Decoding neural responses to temporal cues for sound localization," eLife **2**, doi: `10.7554/eLife.01312.001`.

Greenberg, D., Monaghan, J. J. M., Dietz, M., Marquardt, T., and McAlpine, D. (**2017**). "Influence of envelope waveform on itd sensitivity of neurons in the auditory midbrain," Journal of Neurophysiology **118**(4), 2358–2370, doi: `10.1152/jn.01048.2015`.

Greenwood, D. D. (**1961**). "Critical bandwidth and the frequency coordinates of the basilar membrane," The Journal of the Acoustical Society of America **33**(10), 1344–1356, doi: `10.1121/1.1908437`.

Grothe, B., Pecka, M., and McAlpine, D. (**2010**). "Mechanisms of sound localization in mammals," Physiological Reviews **90**(3), 983–1012, doi: `10.1152/physrev.00026.2009`.

Jeffress, L. (**1948**). "A place theory of sound localization," Journal of comparative and physiological psychology **41**(1), 35—39, doi: `10.1037/h0061495`.

John, M. S., Dimitrijevic, A., and Picton, T. (**2002**). "Auditory steady-state responses to exponential modulation envelopes," Ear and Hearing **23**, 106–117, doi: `10.1097/00003446-200204000-00004`.

Joris, P. X. (**1996**). "Envelope coding in the lateral superior olive. ii. characteristic delays and comparison with responses in the medial superior olive," Journal of Neurophysiology **76**(4), 2137–2156, doi: `10.1152/jn.1996.76.4.2137`.

Joris, P. X., and van der Heijden, M. (**2019**). "Early binaural hearing: The comparison of temporal differences at the two ears," Annual Review of Neuroscience **42**(1), 433–457, doi: `10.1146/annurev-neuro-080317-061925`.

Joris, P. X., and Yin, T. C. (**1992**). "Responses to amplitude-modulated tones in the auditory nerve of the cat," The Journal of the Acoustical Society of America **91**(1), 215–232, doi: `10.1121/1.402757`.

Joris, P. X., and Yin, T. C. (**1995**). "Envelope coding in the lateral superior olive. I. Sensitivity to interaural time differences," Journal of Neurophysiology **73**(3), 1043–1062, doi: `10.1152/jn.1995.73.3.1043`.

Joris, P. X., and Yin, T. C. T. (**1998**). "Envelope coding in the lateral superior olive. iii. comparison with afferent pathways," Journal of Neurophysiology **79**(1), 253–269, doi: `10.1152/jn.1998.79.1.253`.

Kelvasa, D., and Dietz, M. (**2015**). "Auditory model-based sound direction estimation with bilateral cochlear implants," Trends in Hearing **19**, 2331216515616378, doi: `10.1177/2331216515616378`.

Klein-Hennig, M., Dietz, M., Hohmann, V., and Ewert, S. D. (**2011**). "The influence of different segments of the ongoing envelope on sensitivity to interaural time delays," The Journal of the Acoustical Society of America **129**(6), 3856–3872, doi: `10.1121/1.3585847`.

Köppl, C., and Carr, C. E. (**2008**). "Maps of interaural time difference in the chicken's brainstem nucleus laminaris," Biological Cybernetics **9**, 541–559, doi: `10.1007/s00422-008-0220-6`.

Leibold, C., and Grothe, B. (**2015**). "Sound localization with microsecond precision in mammals: What is it we do not understand?," e-Neuroforum **21**(1), 3–10, doi: `doi:10.1515/s13295-015-0001-3`.

Liberman, M. C. (**1978**). "Auditory-nerve response from cats raised in a low-noise chamber," The Journal of the Acoustical Society of America **63**(2), 442–455, doi: `10.1121/1.381736`.

Lindemann, W. (**1986**). "Extension of a binaural cross-correlation model by contralateral inhibition: I. Simulation of lateralization for stationary signals," The Journal of the Acoustical Society of America **80**(6), 1608–1622, doi: `10.1121/1.394325`.

Miller, C. A., Abbas, P. J., and Robinson, B. (**2001**). "Response properties of the refractory auditory nerve fiber," The Journal of the Association for Research in Otolaryngology **2**, 216–232, doi: `10.1007/s101620010083`.

Monaghan, J. J. M., Bleeck, S., and McAlpine, D. (**2015**). "Sensitivity to envelope interaural time differences at high modulation rates," Trends in Hearing **19**, 2331216515619331, doi: `10.1177/2331216515619331`.

Sanes, D. (**1990**). "An in vitro analysis of sound localization mechanisms in the gerbil lateral superior olive," Journal of Neuroscience **10**(11), 3494–3506, doi: `10.1523/JNEUROSCI.10-11-03494.1990`.

Schaette, R., and McAlpine, D. (**2011**). "Tinnitus with a normal audiogram: Physiological evidence for hidden hearing loss and computational model," **31**(38), 13452–13457, doi: `10.1523/JNEUROSCI.2156-11.2011`.

Stern, R. M., and Shear, G. D. (**1996**). "Lateralization and detection of low-freqeuncy binaural stimuli: Effects of distribution of internal delay," The Journal of the Acoustical Society of America **100**(4), 2278–2288, doi: `10.1121/1.417937`.

Takanen, M., Santala, O., and Pulkki, V. (**2014**). "Visualization of functional count-comparison-based binaural auditory model output," Hearing Research **309**, 147–163, doi: `10.1016/j.heares.2013.10.004`.

Tollin, D. J. (**2003**). "The lateral superior olive: A functional role in sound source localization," The Neuroscientist **9**(2), 127–143, doi: `10.1177/1073858403252228` pMID: 12708617.

Tollin, D. J., and Yin, T. C. T. (**2002**). "The coding of spatial location by single units in the lateral superior olive of the cat. i. spatial receptive fields in azimuth," **22**(4), 1454–1467, doi: `10.1523/JNEUROSCI.22-04-01454.2002`.

Tsuchitani, C. (**1988**). "The inhibition of cat lateral superior olive unit excitatory responses to binaural tone bursts. II. The sustained discharges," Journal of Neurophysiology **59**(1), 184–211, doi: `10.1152/jn.1988.59.1.184`.

van de Par, S., and Kohlrausch, A. (**1997**). "A new approach to comparing binaural masking level differences at low and high frequencies," The Journal of the Acoustical Society of America **101**(3), 1671–1680, doi: `10.1121/1.418151`.

Wilson, R., and Collins, A. (**2019**). "Ten simple rules for the computational modeling of behavioral data," eLife **9**, doi: `10.7554/eLife.49547`.

Yin, T. C., Smith, P. H., and Joris, P. X. (**2019**). *Neural Mechanisms of Binaural Processing in the Auditory Brainstem*, 1503–1575 (John Wiley & Sons, Ltd), doi: `10.1002/cphy.c180036`.

Zilany, M. S. A., and Bruce, I. C. (**2006**). "Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery," The Journal of the Acoustical Society of America **120**(3), 1446–1466, doi: `10.1121/1.2225512`.

Zilany, M. S. A., Bruce, I. C., Nelson, P. C., and Carney, L. H. (**2009**). "A phenomenological model of the synapse between the inner hair cell and auditory nerve: Long-term adaptation with power-law dynamics," The Journal of the Acoustical Society of America **126**(5), 2390–2412, doi: `10.1121/1.3238250`.

4. Characterization of the decline in auditory nerve phase locking at high frequencies

4.1. Abstract

The frequency dependence of phase locking in the auditory nerve influences various auditory coding mechanisms. The decline of phase locking with increasing frequency is commonly described by a low-pass filter. This study compares fitted low-pass filter parameters with the actual rate of phase locking decline. The decline is identical across studies and only $40\,\text{dB}$ per decade, corresponding to the asymptotic decline of a second order filter.

4.2. Introduction

The central prerequisite for encoding temporal fine structure (TFS) in the auditory system is the phase locking of auditory nerve (AN) fibers to the frequency of a stimulus (Moore, 2008). This phase-locking is usually quantified by vector strength (VS) (Goldberg and Brown, 1969). The detailed behavior of vector strength across frequency in humans remains unclear and estimates of a so called 'upper frequency limit' vary substantially from 1.5 to 10 kHz (Verschooten *et al.*, 2019). Establishing the vector strength of phase locking in the AN as a function of frequency would be highly informative for elucidating various auditory coding mechanisms from pitch perception to binaural hearing.

It is well known that vector strength in single fibers of the AN declines above a certain frequency (Rose *et al.*, 1967; Johnson, 1980; Joris *et al.*, 1994) that varies across species (Weiss and Rose, 1988a). The frequency above which vector strength becomes insignificant is often referred to as the upper frequency limit (Palmer and Russell, 1986; Verschooten *et al.*, 2015; Joris and Verschooten, 2013). Weiss and Rose (1988a) criticized the comparison of the upper frequency limit among different species stating: "This metric obviously depends upon the method of detection and is generally the highest frequency for which the experimenter detected synchronization in the measurements." The upper frequency limit is therefore ill-defined. Any attempt to define the upper limit for this continually decreasing function depends entirely on the noise floor (Michael Heinz in Verschooten *et al.*, 2019). Despite these shortcomings, the ill-defined characterization of the upper frequency limit is still often used in the auditory literature (Liu *et al.*, 2006; Verschooten *et al.*, 2015; Joris and Verschooten, 2013; Verschooten *et al.*, 2019).

To avoid the above-mentioned shortcomings and to provide a clear definition, Weiss and Rose (1988a) described the frequency dependence using a low-pass filter cascade with the frequency response:

$$G(f) = \frac{g}{\left(\sqrt{1 + \left(\frac{f}{f_0}\right)^2}\right)^n}, \tag{4.1}$$

where

$$f_0 = \frac{f_{\text{corner}}}{\sqrt{2^{\frac{1}{n}} - 1}}, \tag{4.2}$$

here, f is the frequency, f_{corner} the corner frequency, n the order and g the gain at $f = 0$ Hz. In this filter design, a first-order filter is applied n-times. Therefore, the -3 dB point of this filter (f_0) needs to be adjusted by Eq. (4.2) to guarantee that the -3 dB point of the filter cascade is at f_{corner} independent of n.

To estimate the filter order, Weiss and Rose (1988a) compensated for the different corner frequencies across species and fitted a regression line to the five highest frequency points (two from Johnson (1980) and three from Weiss and Rose (1988b)), resulting in a decline of $\approx$ 106 dB/decade. This line is reproduced as the dotted black line in Fig. 4.1. Overall, this decline was related to a filter order of 4-6 by Weiss and Rose (1988a). However, using a cascaded low-pass filter of order $n = 5$ with a corner frequency $f_{corner} = 2500\,\text{Hz}$ does not account for the vector strength, e.g., of the cat data from Johnson (1980). Figure 4.1 clearly shows that the transition band of the filter is too wide and the nominal decline of 100 dB/decade is reached only at much higher frequencies.
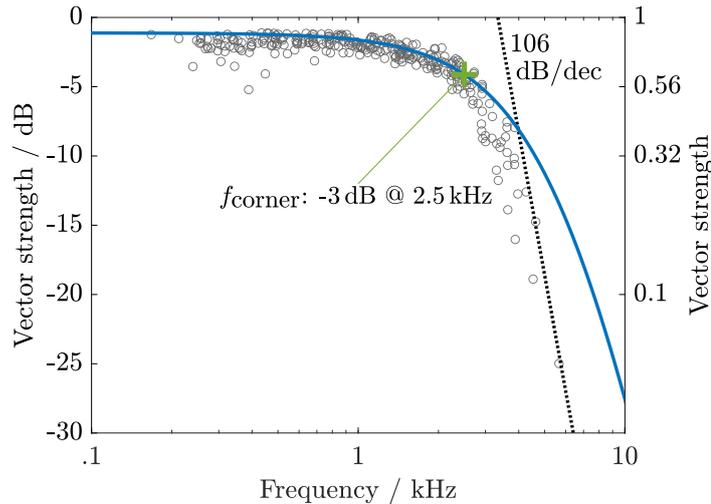
Figure 4.1.: Maximum vector strength for tones at the center frequency of the measured AN fiber as a function of frequency. Circles are data from the cat (Johnson, 1980). Corner frequency (green +) and decline (dotted black line) were estimated by Weiss and Rose (1988a). The transition band of the *cascade* low-pass filter (see Eq. (4.1)) with $f_{corner} = 2500\,\text{Hz}$ and $n = 5$ (solid blue line) is too wide to account for the data.

Despite these discrepancies to physiological data, the filter is prevalent in computational auditory models (see Table. 4.1 for an overview). Some of the studies (e.g., Bernstein and Trahiotis, 1996) motivate their filter order based on the nominal order suggested by Weiss and Rose (1988a), arguably requiring a very low corner frequency of 425 Hz to obtain a sufficiently steep roll-off at higher frequencies. Others, e.g., Heinz *et al.* (2001), instead use a higher filter order for the effect of the inner hair cell (IHC), whereby the combination with their other stages (outer hair cell, synapse) reproduces the phase-locking roll-off from Johnson (1980).

The two goals of the current study were therefore: (1) To revisit the fitting of a low-pass filter to vector strength data, and (2) to quantify the decline of AN phase locking as a function of frequency.

Table 4.1.: Low-pass filter contributing to the decline of AN phase locking as a function of frequency reported by several publications.

	order (n)	corner frequency (f_{corner})[f_0]	type
Weiss and Rose (1988a)	4-6	2500 Hz (cat)	cascade
Lopez-Poveda and Eustaquio-Martín (2006)	2		
Verschooten *et al.* (2015)	(3)		decline 60 dB/dec.
Peterson and Heil (2020)	3	depends on best frequency	Butterworth
Bernstein and Trahiotis (1996)	4	425 Hz	cascade
Breebaart *et al.* (2001a)	5	770 Hz	cascade
Heinz *et al.* (2001)[a]	7	≈ 1500 Hz[b] [4800 Hz]	cascade
Zhang *et al.* (2001)[a]	7	≈ 1200 Hz [3800 Hz]	cascade
Zilany *et al.* (2009)[a]	7	≈ 1000 Hz [3000 Hz][c]	cascade

[a] This filter is applied for only the effect of the inner hair cell (IHC).

[b] According to Eq. (4.2). Heinz *et al.* (2001) reported $f_{\mathrm{corner}} \approx 2500$ Hz for the AN phase locking including their synapse stage.

[c] Used in following versions up to Bruce *et al.* (2018)

4.3. Methods

To determine the decline of AN phase locking with increasing frequency, we propose three functions which we fit to three data sets each: [1] The *cascade* low-pass filter in Eq. (4.1) introduced by Weiss and Rose (1988a), in the following referred to as CASCADE. [2] A Butterworth low-
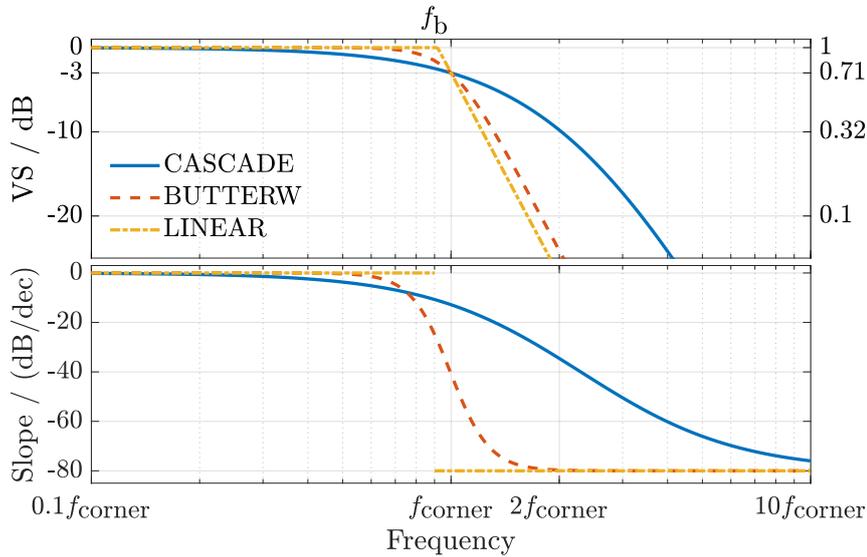
Figure 4.2.: Different filter functions to be fitted to the data with $g = 1$, $n = 4$. CASCADE is from Eq. (4.1), BUTTERW from Eq. (4.3) and LINEAR based on Eq. (4.4). The top panel shows the filter gain or VS across frequency. The lower panel shows the slope as a function of frequency.

pass filter, in the following referred to as BUTTERW, which has a fast transition from passband

to the true decay of the filter order. In this case, the frequency response is given by:

$$G(f) = \frac{g}{\left(\sqrt{1 + \left(\frac{f}{f_{\text{corner}}} \right)^{2n}} \right)}.$$ (4.3)

Note that the only difference to the frequency response of CASCADE is that the exponent/ order n is not to the power of the whole denominator as in Eq. (4.1). [3] A constant decline archived by fitting a broken-stick function to the data in double logarithmic space:

$$\hat{G}(f) \ /\text{dB} = 20 \ \log_{10}(G(f)) = \begin{cases} m \ \log_{10}(f_{\text{b}}/f) + g & , \text{for } f \geq f_{\text{b}} \\ g & , \text{for } f < f_{\text{b}}, \end{cases}$$ (4.4)

where g determines the gain at low frequencies and m the slope of the function after the knee point f_{b}. The resulting function is in the following referred to as LINEAR and is plotted with the other functions in Fig. 4.2. LINEAR is also defined by f_{corner} the -3 dB point by adjusting the knee point f_{b}.

As data sets, we used vector strength derived from single-unit neuron recordings: from AN fibers in Johnson (1980) from cats, as well as from AN fibers in Heeringa *et al.* (2020) from gerbils and in Palmer and Russell (1986) from guinea pigs. As the error metric, we chose the root mean square error (RMSE) in the logarithmic space. The RMSE for all three functions describing the vector strength in the three data sets was calculated as a function of order n, gain g and corner frequency f_{corner}. The minimum RMSE fit was derived by a grid search, the range and step size of the individual parameters are shown in Table 4.2. For the calculations, MATLAB (MathWorks, Natick, MA) was used. By definition, the filter order is an integer number but the filter equations produce meaningful output also for any positive non-integer number n. As the focus of the present study is on describing the data, some non-integer values for n were included.

Table 4.2.: Range and step size for the grid search for the minimum RMSE fit.

	range	step
order n	1 - 16	1^a
gain g	0.6 - 1	0.01
corner frequency f_{corner}	1000-3000 Hz	10 Hz

[a]Smaller if necessary.

4.4. Results
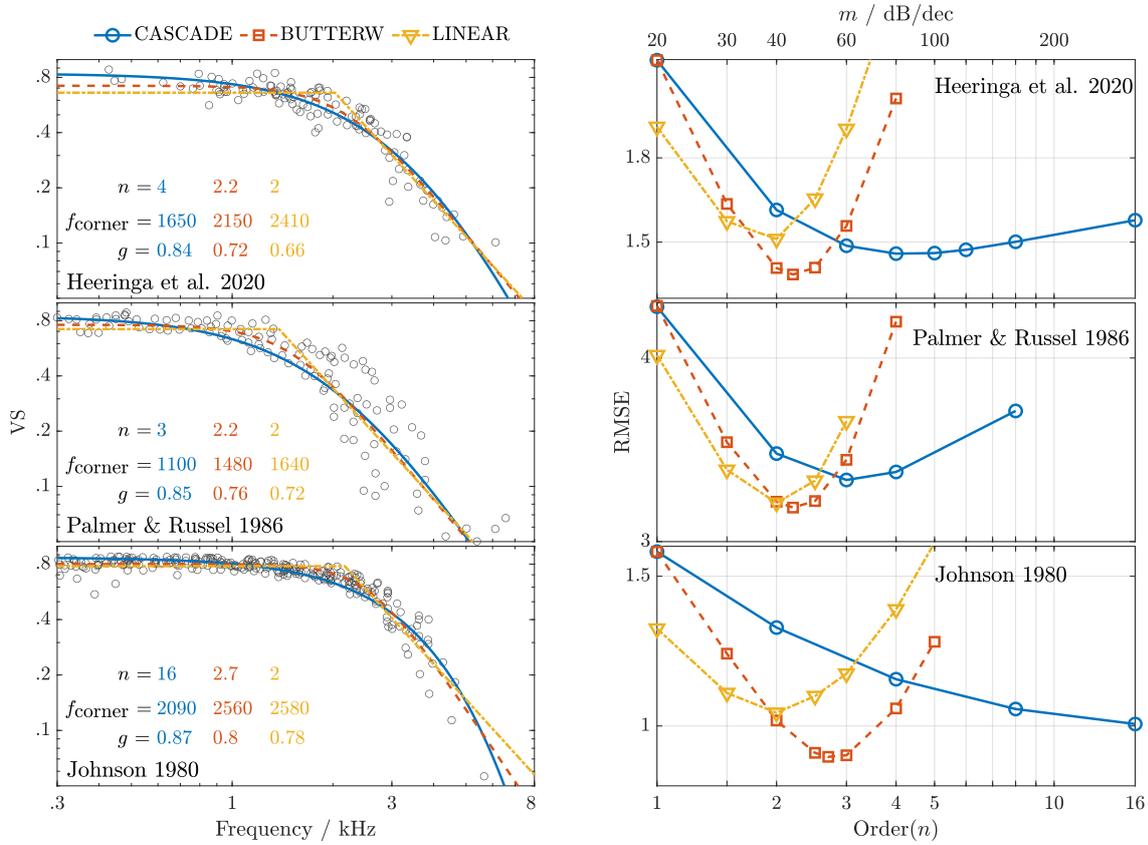
Figure 4.3.: (**left**) The three panels show the vector strength for different data sets and the minimum RMSE solution for the three filter functions (color code). (**right**) RMSE across filter order n for the three different functions (color code) and three different data sets.

For eight of the nine fitted functions (3 types times 3 data sets), a global best fit was obtained within the above-mentioned parameter range, i.e. not at the lowest or highest parameter values. The only exception is CASCADE, which showed no local minimum for the order in Johnson (1980). The best solutions are shown above the respective data sets in the left column in Fig. 4.3. The right column of Fig. 4.3 shows the RMSE across order n for the three functions and three data sets. LINEAR results in the lowest decline estimates followed by BUTTERW, followed by CASCADE. For CASCADE the best fit low-pass filter order is 3 for Palmer and Russell (1986) and 4 for (Heeringa *et al.*, 2020). For these two data sets the roll-off within the shown range is about 40 dB/dec, as the *cascade* filters have not reached their maximum roll-off. For CASCADE, the RMSE declines across n for the Johnson (1980) data set and shows no local minimum for $n \leq 16$. This decrease is caused by the data point with the highest frequency and lowest vector strength. For all data sets BUTTERW provides the lowest RMSE solutions. Fits were similar when only data with frequencies > 1.8 kHz or > 3 kHz were considered to estimate the roll off.

4.5. Discussion

The decline in vector strength at high frequencies is much flatter than implied by the nominal low-pass filter order. Especially for the *cascade* filter, the slope at intermediate frequencies is different from the asymptotic values (cf. Fig 4.2). Linear fits of the high-frequency roll-off suggest a decline of about 40 dB/dec, corresponding to the asymptotic decline of a second order low-pass filter.

As the low-pass characteristic of the phase locking decline is the outcome of a sequential interaction of several processing stages, the use of a *cascade* filter is a logical means for a biologically feasible simulation of the underlying processes. The low-pass filter order of this type is at least 3 and even much larger orders provide good fits. The roll-off within the range shown is generally no more than about 60 dB/dec, still considerably lower than the 106 dB/dec derived by Weiss and Rose (1988a). While it produced generally accurate fits, the biologically less plausible filters had lower RMS errors in all three data sets.

While the corner frequency appears to differ between different species, Palmer and Russell (1986) and Weiss and Rose (1988a) concluded that there is a similar decline across species. The similarity in decline across species is unsurprising, considering that the free parameters of the resistor-capacitor (RC) low-pass filter can only influence the corner frequency but not the filter order (Lopez-Poveda and Eustaquio-Martín, 2006). On the other hand, Altoè *et al.* (2018) showed that the voltage-dependent activation of the K^+ channels in the IHC enhances the phase-locking properties within a small resonant frequency range. The decline above the resonant frequency range is consequently steeper than with a conventional RC low-pass filter.

Without investigating specific mechanisms in the IHC-AN complex, this study only analyses the function of VS with respect to frequency that can be used in functional models. Thus, we do not claim that any of the filters we employ correspond to the physiological processes. From the fitted parameter (see Figure 4.2), a small difference in slope can be deduced, with the corner frequency varying between species, as was reported by Weiss and Rose (1988a). This finding suggests that the common practice of using the same filter order in models of the human auditory system as determined for animal models (e.g. Bernstein and Trahiotis, 1996) is plausible.

The decline of vector strength across frequency is particularly important for understanding the interaural phase difference (IPD) encoding mechanism. Phase-locking information is a strict prerequisite, as it is the only available cue for IPD encoding. In normal hearing humans and for pure tones, it is well established that the sensitivity to IPD decreases rapidly at frequencies above ≈ 1300 Hz (Brughera *et al.*, 2013; Klug and Dietz, 2022). A common hypothesis is that this decline in IPD sensitivity is caused by the decline in phase locking of the AN fibers (Joris and Verschooten, 2013; Verschooten *et al.*, 2019). In computational models of human IPD perception, the decline of AN phase locking usually contributes considerably, or even exclusively, to the decline in IPD sensitivity (Breebaart *et al.*, 2001b; Brughera *et al.*, 2013; Bouse *et al.*, 2019). However, the decline in vector strength in the AN ($m \leq 60$ dB/dec) is not sufficient to account for the dramatic decline of IPD sensitivity in humans occurring in the narrow range between 1300 and 1500 Hz ($m \geq 150$ dB/dec), as reported in Klug and Dietz (2022). Therefore, the high-frequency limit of IPD sensitivity is a less suitable correlate for the AN phase locking 'limit' than previously thought (Verschooten *et al.*, 2019; Bernstein and Trahiotis, 1996) and another effect must contribute to limiting IPD sensitivity.

4.6. Conclusion

The nominal decline of 80 to 120 dB/dec reported by Weiss and Rose (1988a) does not accurately describe the decline of vector strength in AN fibers across frequency. In the biophysically plausible and widely used *cascade* filter, the slope changes slowly with frequency, so the slope at intermediate frequencies differs substantially from its asymptote. The relevant vector strength values decline with approximately 40 dB/dec - the asymptotic decline of a second order filter.

4.7. References

Altoè, A., Pulkki, V., and Verhulst, S. (**2018**). "The effects of the activation of the inner-hair-cell basolateral K+ channels on auditory nerve responses," Hearing Research **364**, 68–80, doi: `10.1016/j.heares.2018.03.029`.

Bernstein, L. R., and Trahiotis, C. (**1996**). "The normalized correlation: Accounting for binaural detection across center frequency," The Journal of the Acoustical Society of America **100**(6), 3774–3784, doi: `10.1121/1.417237`.

Bouse, J., Vencovský, V., Rund, F., and Marsalek, P. (**2019**). "Functional rate-code models of the auditory brainstem for predicting lateralization and discrimination data of human binaural perception," J. Acoust. Soc. Am. **145**(1), 1–15, doi: `10.1121/1.5084264`.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001a**). "Binaural processing model based on contralateral inhibition. i. model structure," J. Acoust. Soc. Am. **110**(2), 1074–1088, doi: `10.1121/1.1383297`.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001b**). "Binaural processing model based on contralateral inhibition. ii. dependence on spectral parameters," J. Acoust. Soc. Am. **110**(2), 1089–1104, doi: `10.1121/1.1383298`.

Bruce, I. C., Erfani, Y., and Zilany, M. S. (**2018**). "A phenomenological model of the synapse between the inner hair cell and auditory nerve: Implications of limited neurotransmitter release sites," Hearing Research **360**, 40–54, doi: `10.1016/j.heares.2017.12.016` computational models of the auditory system.

Brughera, A., Dunai, L., and Hartmann, W. M. (**2013**). "Human interaural time difference thresholds for sine tones: The high-frequency limit," J. Acoust. Soc. Am. **133**(5), 2839–2855, doi: `10.1121/1.4795778`.

Goldberg, J. M., and Brown, P. B. (**1969**). "Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization.," Journal of Neurophysiology **32**(4), 613–636, doi: `10.1152/jn.1969.32.4.613`.

Heeringa, A. N., Zhang, L., Ashida, G., Beutelmann, R., Steenken, F., and Köppl, C. (**2020**). "Temporal coding of single auditory nerve fibers is not degraded in aging gerbils," Journal of Neuroscience **40**(2), 343–354, doi: `10.1523/JNEUROSCI.2784-18.2019`.

Heinz, M. G., Colburn, H. S., and Carney, L. H. (**2001**). "Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the Auditory Nerve," Neural Computation **13**(10), 2273–2316, doi: `10.1162/089976601750541804`.

Johnson, D. H. (**1980**). "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," The Journal of the Acoustical Society of America **68**(4), 1115–1122, doi: `10.1121/1.384982`.

Joris, P. X., Carney, L. H., Smith, P. H., and Yin, T. C. (**1994**). "Enhancement of neural synchronization in the anteroventral cochlear nucleus. i. responses to tones at the characteristic frequency," Journal of Neurophysiology **71**(3), 1022–1036, doi: `10.1152/jn.1994.71.3.1022`.

Joris, P. X., and Verschooten, E. (**2013**). "On the limit of neural phase locking to fine structure in humans," in *Basic Aspects of Hearing*, edited by B. C. J. Moore, R. D. Patterson, I. M. Winter, R. P. Carlyon, and H. E. Gockel, Springer New York, New York, NY, pp. 101–108, doi: `10.1007/978-1-4614-1590-9_12`.

Klug, J., and Dietz, M. (**2022**). "Frequency dependence of sensitivity to interaural phase differences in pure tones," The Journal of the Acoustical Society of America **152**(6), 3130–3141, doi: `10.1121/10.0015246`.

Liu, L.-F., Palmer, A. R., and Wallace, M. N. (**2006**). "Phase-locked responses to pure tones in the inferior colliculus," Journal of Neurophysiology **95**(3), 1926–1935, doi: `10.1152/jn.00497.2005`.

Lopez-Poveda, E. A., and Eustaquio-Martín, A. (**2006**). "A biophysical model of the inner hair cell: The contribution of potassium currents to peripheral auditory compression," Journal of the Association for Research in Otolaryngology **7**, 218–235, doi: `10.1007/s10162-006-0037-8`.

Moore, B. J. C. (**2008**). "The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people," Journal of the Association for Research in Otolaryngology **9**, 399–406, doi: `10.1007/s10162-008-0143-x`.

Palmer, A. R., and Russell, I. J. (**1986**). "Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells," Hearing Research **24**(1), 1–15, doi: `10.1016/0378-5955(86)90002-X`.

Peterson, A. J., and Heil, P. (**2020**). "Phase locking of auditory nerve fibers: The role of lowpass filtering by hair cells," Journal of Neuroscience **40**(24), 4700–4714, doi: `10.1523/JNEUROSCI.2269-19.2020`.

Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (**1967**). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey.," Journal of Neurophysiology **30**(4), 769–793, doi: `10.1152/jn.1967.30.4.769`.

Verschooten, E., Robles, L., and Joris, P. X. (**2015**). "Assessment of the limits of neural phase-locking using mass potentials," Journal of Neuroscience **35**(5), 2255–2268, doi: `10.1523/JNEUROSCI.2979-14.2015`.

Verschooten, E., Shamma, S., Oxenham, A., Moore, B., Joris, P., Heinz, M., and Plack, C. (**2019**). "The upper frequency limit for the use of phase locking to code temporal fine structure

in humans: A compilation of viewpoints," Hearing Research **377**, doi: `10.1016/j.heares.2019.03.011`.

Weiss, T., and Rose, C. (**1988**a). "A comparison of synchronization filters in different auditory receptor organs," Hearing Research **33**(2), 175–179, doi: `10.1016/0378-5955(88)90030-5`.

Weiss, T., and Rose, C. (**1988**b). "Stages of degradation of timing information in the cochlea: A comparison of hair-cell and nerve-fiber responses in the alligator lizard," Hearing Research **33**(2), 167–174, doi: `10.1016/0378-5955(88)90029-9`.

Zhang, X., Heinz, M. G., Bruce, I. C., and Carney, L. H. (**2001**). "A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression," The Journal of the Acoustical Society of America **109**(2), 648–670, doi: `10.1121/1.1336503`.

Zilany, M. S. A., Bruce, I. C., Nelson, P. C., and Carney, L. H. (**2009**). "A phenomenological model of the synapse between the inner hair cell and auditory nerve: Long-term adaptation with power-law dynamics," The Journal of the Acoustical Society of America **126**(5), 2390–2412, doi: `10.1121/1.3238250`.

5. Frequency dependence of sensitivity to interaural phase differences in pure tones

5.1. Abstract

It is well established that in normal-hearing humans, the threshold of interaural time differences (ITDs) for pure tones increases dramatically above about 1300 Hz, only to become unmeasurable above 1400 Hz. However, physiological data and auditory models suggest that the actual decline in sensitivity is more gradual and only appears to be abrupt because the maximum of the psychometric function dips below the threshold proportion correct, e.g. 0.794. Published data only report thresholds at certain proportions correct but not the decline of proportions correct or of the sensitivity index d' with increasing frequencies. Here, we present pure-tone behavioral data obtained with a constant stimulus procedure. Seven of nine subjects showed proportions correct above 0.9 at 1300 Hz and virtually no sensitivity at 1500 Hz (proportion correct within 0.07 of chance level). This corresponds to a sensitivity decline of 46-78 dB/oct, much steeper than predicted by existing models or by the decline of phase-locking of the auditory nerve fibers in animal data.

5.2. Introduction

Binaural cues provide humans with the ability to discriminate azimuthal angles of sound direction. The cue that enables the discrimination of angles as small as 1° is the interaural time difference (ITD) in the temporal fine structure (TFS) (Mills, 1958), abbreviated here as ITD_{TFS}. To measure human sensitivity to ITD_{TFS} in isolation, pure tone presentation via headphones is necessary. Usually, two intervals are used with the same magnitude of ITD but with opposite signs (i.e. $\text{ITD}_2 = \text{-ITD}_1$; Brughera *et al.*, 2013). Sensitivity is then often reported as the ΔITD threshold:

$$\Delta\text{ITD} = \text{ITD}_1 - \text{ITD}_2 = 2\,\text{ITD}, \tag{5.1}$$

with $\text{ITD} = |\text{ITD}_1| = |\text{ITD}_2|$, i.e. the difference in ITD between two successive sounds at which a certain proportion correct (e.g., $p_{\text{c}} = 0.794$) is achieved. Previous studies have underlined the remarkable ITD_{TFS} sensitivity of trained normal-hearing subjects. The maximum sensitivity, i.e. the lowest ΔITD thresholds for pure-tone sensitivity, have consistently been reported for frequencies of 700-1000 Hz, where the thresholds can be as low as 10-30 µs (Brughera *et al.*, 2013; Zwislocki and Feldman, 1956).

However, despite the remarkably low thresholds at low frequencies, ITD sensitivity deteriorates rapidly for frequencies above 1 kHz and $p_{\text{c}} = 0.794$ thresholds become unmeasurable above 1.4 kHz (Brughera *et al.*, 2013). This sudden increase in ΔITD thresholds is called upper frequency limit of ITD_{TFS} sensitivity. The thresholds only provide a single data point for the unknown psychometric function $p(f, \Delta\text{ITD})$. Further, there is no information about sensitivity above 1.4 kHz.

The upper frequency limit of ITD_{TFS} perception has been related to the upper limit of phase-locking of auditory nerve (AN) fibers to the TFS in humans (Verschooten *et al.*, 2019). Synchrony index (SI) is the most-used metric for phase-locking, also referred to as vector strength (VS) (Goldberg and Brown, 1969). Phase-locking data of AN fibers are only available for non-human animals, but interesting common features emerge when comparing data from different mammalian species (Kiang *et al.*, 1965; Rose *et al.*, 1967; Johnson, 1980). The decline of SI across frequency has a very similar shape and steepness across species, but the corner frequency varies when expressed as a low-pass filter (Weiss and Rose, 1988a). For all mammalian species for which data is available, there is still a highly significant phase-locking at 1.4 kHz which, at higher frequencies, declines at approximately ≤18 dB/oct (Weiss and Rose, 1988a,b; Verschooten *et al.*, 2015; Klug *et al.*, 2023). Phenomenological models of human ITD perception commonly explicitly include such a low-pass filter, which can correctly simulate the reduced binaural unmasking with increasing frequency (Breebaart *et al.*, 2001b). However, such filters do not accurately simulate the abrupt increase of pure tone threshold ITDs (Breebaart *et al.*, 2001b; Bouse *et al.*, 2019).

The discrepancy between the models and experimental data (Brughera *et al.*, 2013) is not particularly striking below 1.4 kHz. But the models can, of course, provide simulations also at higher frequencies: Bouse *et al.* (2019) simulate a $p_{\text{c}} = 0.794$ threshold of 160 µs at 1.45 kHz. Using a lower $p_{\text{c}} = 0.707$, Breebaart *et al.* (2001b) are even able to simulate a threshold of 50 µs at 1.5 kHz. However, as no experimental data exists, it is unclear if the maximum possible human p_{c} at 1.45 kHz and 1.5 kHz is just marginally lower or far below the threshold proportions of 0.794 or 0.707, respectively. To determine the ITD sensitivity roll-off experimentally, the p_{c}

must be measured directly to estimate how it changes over frequency for a fixed ΔITD. In Section 5.3 basic concepts of signal detection theory are reviewed and used to propose a basic model of the psychometric function.

The goals of the present study are [1] to provide the currently missing psychophysical data on ITD$_{\text{TFS}}$ sensitivity in the frequency range between 1300 and 1500 Hz, [2] to extract and describe the decline of ITD$_{\text{TFS}}$ sensitivity in terms of the sensitivity index (d',"d prime"), with a focus on the 1300-1500 Hz frequency range, and [3] to relate existing models and concepts to the low-pass characteristic.

5.3. Model of the psychometric function

In signal detection theory, the sensitivity of sensory systems in different psychophysical tasks is often quantified with the sensitivity index d' (Green and Swets, 1966). A principal assumption thereby is that the various sensory events (triggered by a stimulus) can be mapped to a single dimension of experimental interest. Each observation (of the target or the reference) produces one event on this dimension. Consequently, the stimuli are represented as two partially overlapping distributions. For ITD discrimination in a two-alternative forced-choice (2AFC) paradigm, it is assumed that the subject compares the internal laterality representation of the first interval to the representation of the second interval. d' is considered as the ideal observer performance

$$d' = \frac{\mu_1 - \mu_2}{\sigma} = \frac{\Delta\mu}{\sigma}, \tag{5.2}$$

where μ_1 and μ_2 are the means of the distributions of the internal representation and σ is the standard deviation of each distribution (assuming both standard deviations are reasonably similar). Further, the distributions are assumed to be Gaussian, so the discrimination index can be transformed to a discrimination probability p in a 2AFC experiment by

$$p = \Phi\left(\frac{d'}{\sqrt{2}}\right), \tag{5.3}$$

with Φ the normal cumulative distribution function. This transformation is done implicitly throughout the text.

Psychophysically it cannot be distinguished if a change in sensitivity is caused by a change of $\Delta\mu$ or by a change of σ of the internal representation. That, however, provides the freedom to assign all acoustic differences between target and reference waveform to the numerator $\Delta\mu$ and all differences caused by the processing of the auditory system to the denominator σ. Regarding the acoustical difference, the two intervals differ in their interaural phase difference (IPD) and the magnitude of the difference is ΔIPD. The maximum physical difference in phase space is 180° or π radians. In the present case with IPD$_1$ = $-$IPD$_2$, the maximum (π) difference for ΔIPD occurs for a nominal IPD of $\pi/2$ radians only to become zero again at an IPD of π. The acoustical difference $\Delta\mu$ can thus be expressed as

$$\Delta\mu := \sin(\text{IPD}). \tag{5.4}$$

With this definition, the remaining term $1/\sigma$ becomes a measure of observer sensitivity and Eq.

(5.2) can be rewritten as

$$d'(f, \Delta\mathrm{IPD}) = \frac{\sin(\mathrm{IPD})}{\sigma(f, \mathrm{IPD})}. \tag{5.5}$$

The focus of this study is describing d', or better the observer sensitivity $1/\sigma$, as a function of frequency. To compare this function to the decline of phase-locking and to the low-pass filter of the models, the final quantity of interest is the decline of $1/\sigma(f)$. To compare the decline across different modalities and units (neural synchrony index, vs. model filter, vs. psychophysically obtained sensitivity index) changes of all quantities are expressed in dB. The resulting decline unit is dB/oct and thus expresses the decline while the frequency doubles. Therefore, a $\log_2$-frequency transformation is necessary to obtain decline m from linearly fitting d':

$$10 \log_{10}(d'(f)) = -m \log_2(f/\mathrm{Hz}) + b. \tag{5.6}$$

We calculated the least squares solution for m, with b as a second free parameter determining the y-axis intercept in this linear fit, which is not used for further analysis.

The decline m in Brughera *et al.* (2013) behavioral data can be predicted by combining Eq. (5.5) and Eq. (5.6), see Appendix 5.9.2. When applied to the data of Brughera *et al.* (2013), the decline estimates from this data set, range from $30\,\mathrm{dB/oct}$ to $76\,\mathrm{dB/oct}$, depending on frequency range and subject. Thus, a further motivation for this study is to test whether or not Eq. (5.5) provides a valid basis for the psychometric function and if the estimation of the decline is confirmed by Experiment II in Sec. 5.6.

5.4. Experimental Methods

5.4.1. Subjects

Subjects were limited to young adults with audiometric thresholds equal to or less than $15\,\mathrm{dB\,HL}$ at octave-spaced frequencies from 125 to $6000\,\mathrm{Hz}$. The interaural asymmetry in audiometric threshold was less or equal to $10\,\mathrm{dB}$ at all of the frequencies tested. Since Bernstein and Trahiotis (2016) found a significant dependence of dichotic detection thresholds on the audiometric thresholds at $4\,\mathrm{kHz}$, a maximum audiometric threshold of $5\,\mathrm{dB\,HL}$ at this frequency was an additional criterion for the subjects. A total of 9 normal-hearing trained subjects aged between 21 and 29 years (average age $= 27$, F $= 3$, M $= 6$) participated in the experiment.

The total duration of the measurements was 15-20 hours per subject, performed in a different number of sessions for each subject. Within a session (45 minutes to 2 hours), subjects could take as many breaks as they wanted. Two of the nine subjects were lab members and were well informed about the purpose and methods of the study. They did not receive any extra compensation while the other subjects were paid. The study was approved by the Ethics committee of the University of Oldenburg.

5.4.2. Procedure

The procedure was a two-interval, two-alternative forced-choice task (2I-2AFC). The subject was required to respond according to whether the tone in the second interval was perceived to the left or the right of that presented in the first interval. The stimuli had synchronous onset and offset gating in both ears, so that the tones differed only in their IPD, i.e. their ongoing ITD_{TFS}.

Tones were presented to the subject via Sennheiser HD-650 headphones at a level of 70 dB SPL. Levels where measured with a sound-level meter and an Artificial Ear Type 4153 (Brüel & Kjær). Digital-analog conversion was carried out by ADI-s DAC FS (RME) with 32 bit and a 48 kHz-sampling rate. Stimuli were generated digitally using the AFC-software package (Ewert, 2013). The subjects were seated in a double-walled, sound-attenuating booth and responded by pressing a key on a standard computer keyboard. Visual feedback was provided after each trial. The 300-ms tone duration included 50-ms $\cos^2$ rise-decay ramps. A 50-ms silent interval separated the two intervals. A next pair of intervals was presented 500 ms after the subject responded.

The interaural differences of the stimuli presented in the two intervals were symmetrical around zero (Hafter *et al.*, 1979; Henning, 1983; Brughera *et al.*, 2013; Thavam and Dietz, 2019), so that in one of the two intervals, the right ear was leading, in the other the left ear led by the same IPD. Subjects could thus make their decision based on a ΔIPD difference between the two intervals.

While the focus of this study was on the frequency range ≥ 1300 Hz, Sec.5.5 will elaborate on Experiment I, where we performed measurements in the frequency range from 250 to 1200 Hz. In Experiment II, we address the primary aim of this study - to derive the sensitivity decline across frequency - by measuring with constant ΔIPDs in the frequency range of 1300 to 1500 Hz. To obtain a fine frequency resolution, a 50-Hz spacing was used; details are provided in Sec. 5.6.

Audiograms and results for all subjects are provided in the supplementary material to this paper.

5.5. Experiment I

Sensitivity to low-frequency pure tones is a fundamental aspect of binaural hearing. Therefore, the first purpose of Experiment I was to provide a comprehensive data set for future reference or analysis. Previous studies mostly focused on the frequency dependence of threshold ITDs at a certain p_c value (e.g., Brughera *et al.*, 2013; Zwislocki and Feldman, 1956). We chose an adaptive staircase procedure to collect most responses well above chance and well below ceiling performance. We then display and analyze the proportions correct at all ΔIPDs included during the adaptive runs, not the typically reported threshold ITDs.

Measuring well below the frequency limit means measuring at small IPDs. For reference IPDs from zero to $\pi/4$ radians sensitivity to changes in IPD (i.e. to ΔIPD) is best and reasonably independent of the reference IPD (Yost, 1974). The second purpose of Experiment I was thus to test if $\sigma(f, \text{IPD})$ can be assumed to be IPD-independent up to $\pi/4$ radians, i.e. up to ΔIPD = $\pi/2$ radians. If this is the case, then d' is proportional to $\sin(\text{IPD})$ (see. Eq. 5.5). In other words, for each subject and each frequency the psychometric function $p_c(\Delta\text{IPD})$ can then be fitted with a single free parameter σ.

5.5.1. Measurement Table

In Experiment I, different fixed frequencies (250, 500, 650, 800, 1000, 1100, 1200 Hz) were presented. For each frequency, an adaptive 'three-down, one-up' staircase procedure controlled the ΔIPD, i.e. the ΔIPD was decreased after three correct responses in a row and increased after each incorrect response, asymptoting a p_c of 0.794 (Levitt, 1971). Each adaptive track started at a ΔIPD of $0.23\,\pi$ radians. This starting value was selected based on the results of a pilot experiment, where it resulted in proportions correct well over 0.9. The initial step size was a factor of 2, which was reduced to 1.414 and 1.189 after the first and second 'down-up-reversal'. An adaptive track was terminated after 10 reversals at the minimum step size. Starting at 250 Hz, the tone frequency was increased after each track, up to 1200 Hz, and then decreased after a second track at 1200 Hz. This sequence was repeated eight times, resulting in a total of 16 adaptive tracks per frequency.

5.5.2. Results with fitted psychometric functions

The usual threshold calculation by reversals was omitted. For each combination of frequency and ΔIPD, the p_c was calculated. Although the number of presentations N varied due to the adaptive procedure, it typically exceeded $N = 200$ near the threshold, corresponding to $p_c = 0.794$. This is reflected by the error bars in Fig. 5.1 denoting the 95% confidence intervals (Johnson *et al.*, 1993). The psychometric functions shown as solid lines in Fig. 5.1 indicate a fit based on Eq. (5.5) [d' to p_c and ΔIPD to IPD transformation applied], with the free parameter σ to all responses from the 16 runs for each frequency and subject. Alternatively, a different σ-value can be derived for each data point, i.e. for each ΔIPD. We did not find a trend in these values over the fairly small ΔIPD range, and the IPD-independent σ results in good fits for all frequencies and subjects. The maximum likelihood fit was derived by a grid search with steps of 0.001 from 0.001 to 0.8. Details about the likelihood function and the merge over ΔIPD can be found in the appendix [see Eqs. (5.10) & (5.12)]. The resulting ITD thresholds are shown in Section 5.8.1.
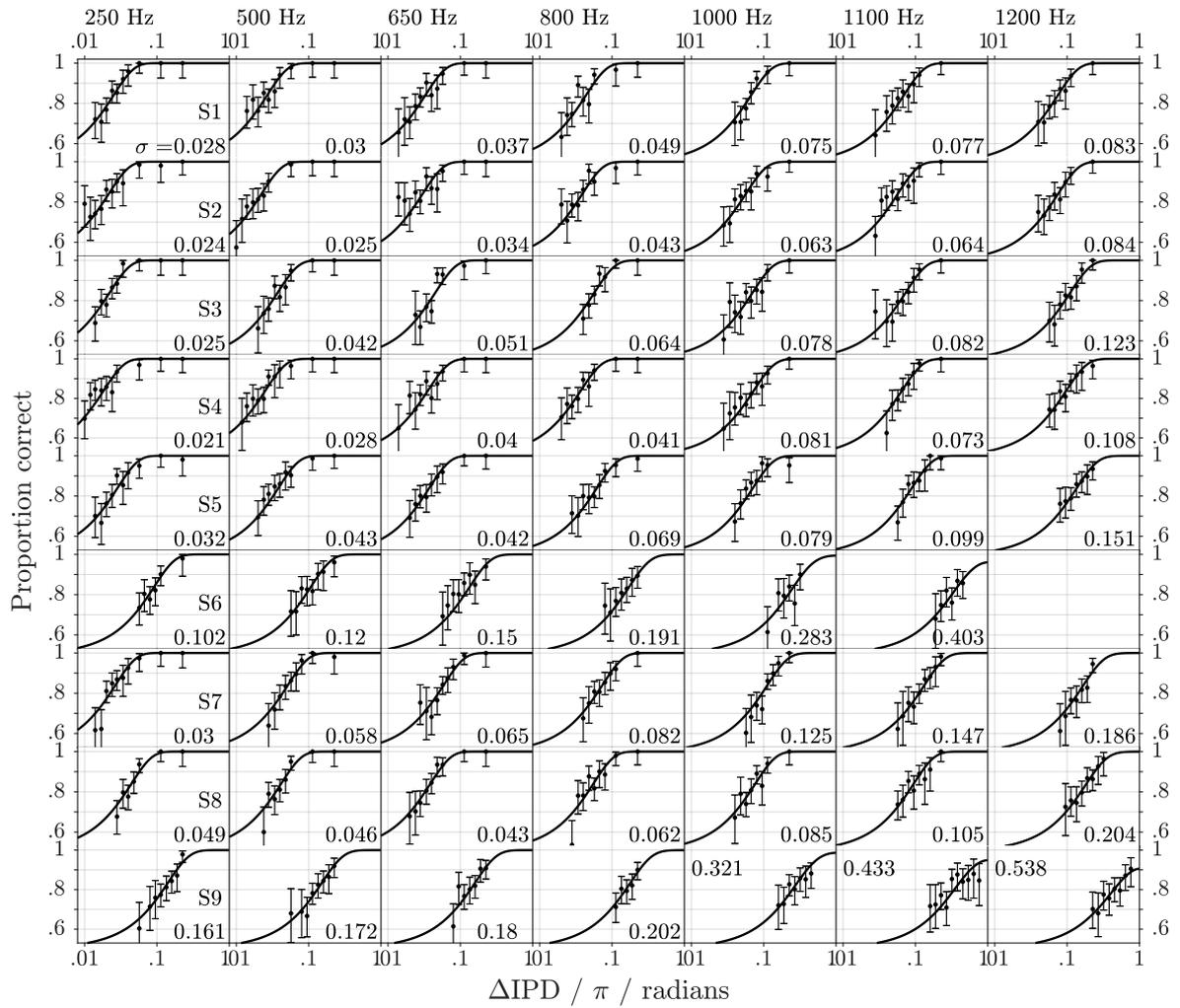
Figure 5.1.: Psychometric functions and fits. The nine rows show proportion correct (p_c) over ΔIPD of the nine subjects (S1-S9). Different frequencies are shown in different columns. The error bars denote the 95 % confidence level derived from the binomial distribution. Therefore, the confidence-interval size depends on the p_c and on the number of times measured (N). Since an adaptive method was used, N varies with subject and condition. Only data from ΔIPDs presented at least 48 times during the 16 adaptive tracks are plotted. The solid lines represent the fit with free parameter σ as given by values in the panel corners.

5.6. Experiment II

5.6.1. Measurement Table

In Experiment II, different fixed frequencies (1300, 1350, 1400, 1450, and 1500 Hz) were presented in combination with fixed ΔIPDs (0.1 π , 0.2 π, 0.4 π, 0.6 π, 0.8 π, 1.0 π, 1.2 π, and 1.4 π, radians). We chose a constant stimulus procedure to get the same number of repetitions for conditions with high and low p_c. The small differences in sensitivity across subjects during the pilot-testing phase led us to decide for this approach. Not all possible combinations were measured: For example, as the performance for $f \geq 1450$ Hz was near chance level in the pilot experiment, we only measured at ΔIPD of 0.4 π rad. and 1 π rad. at those highest frequencies. Each combination was presented in blocks of 50 trials. The frequency was changed upwards from 1300 Hz, then downwards from 1500 Hz after all ΔIPDs had been presented in descending order. In total 56 blocks (28 conditions), were measured four times for each ΔIPD, thus at each frequency resulting in 400 trials per combination. Subject 6 was not able to achieve $p_c > 0.7$ for 1300 Hz, thus the performance was additionally measured for 1200 Hz, but not for 1450 and 1500 Hz.

5.6.2. Results
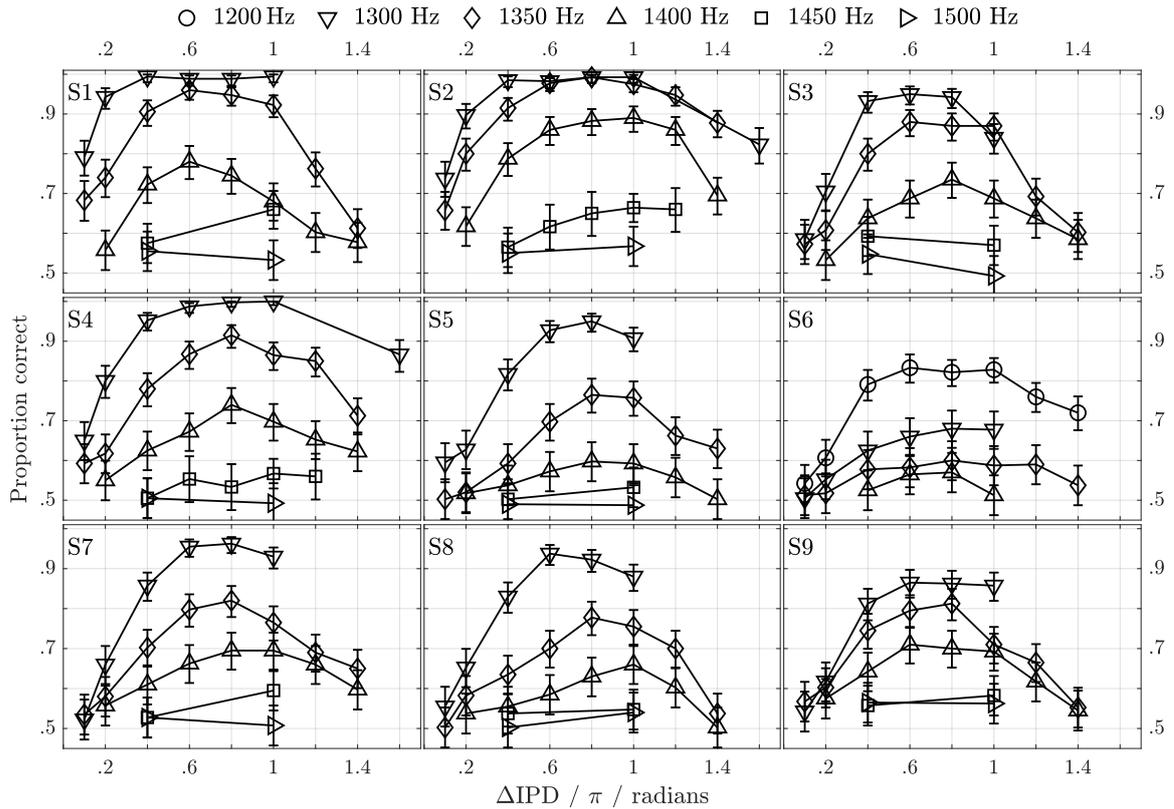
Figure 5.2.: The nine panels portray the proportion correct (p_c) of the nine subjects (S1-S9). Error bars denote the 95 % confidence level. Symbols denote different frequencies as show in the legend on the top. Only S6 has data for 1200 Hz.

Figure 5.2 displays the p_c across ΔIPD for the different frequencies. The panels of the figure each contain data for one of 9 different subjects. Error bars represent the 95 % confidence intervals (Johnson *et al.*, 1993). Three phenomena stand out in the data:

(i) Nearly all subjects (7 out of 9) show performance of $p_c > 0.9$ at 1300 Hz, which drops within a 15% frequency increase (2.5 semitones) to $p_c < 0.6$ at 1500 Hz.

(ii) In general, performance increased with ΔIPD until ΔIPD $= 0.8\pi$ rad., beyond which performance started to decrease.

5.6.3. Effect of IPD on sensitivity

The dependence of the performance on IPD is reflected in the course of sensitivity in Fig. 5.2. Within these data, ΔIPD ranges from $0.1\,\pi$ to $1.4\,\pi$ radians which allows us to analyze $\sigma(f, \text{IPD})$. It is not possible to collect this data with the symmetric paradigm at lower frequencies, due to ceiling performance.
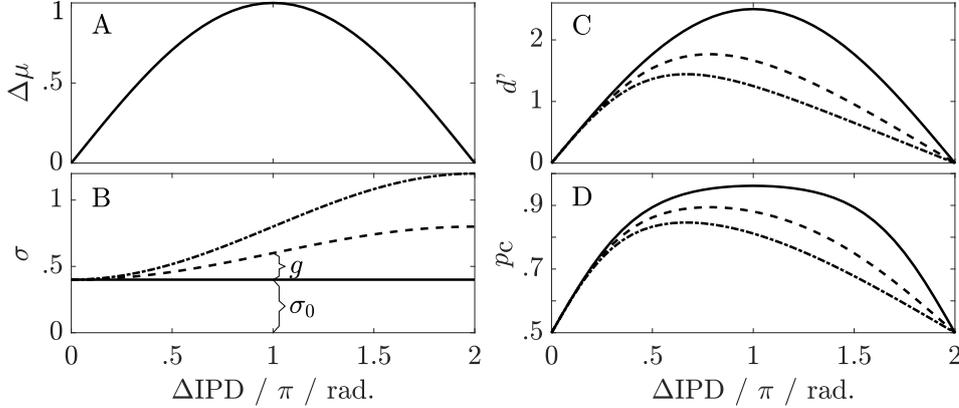
Figure 5.3.: Illustration of the psychometric function. **A**: Stimulus difference between target and reference, cf. Eq. (5.4). **B**: Three examples of Eq. (5.7) using values 0, 0.2, and 0.4 for the parameter g_m, expressing the IPD dependence of the IPD sensitivity. The measure σ is inversely related to IPD sensitivity. **C**: Resulting psychometric functions expressed as sensitivity index $d' = \Delta\mu/\sigma$, cf. Eq. (5.2). **D**: Same as C but transformed into proportion correct (p_c) by Eq. (5.3). The psychometric functions differ from the typical sigmoidal shape because of the linear x-axis and the periodicity of ΔIPD.

The largest stimulus difference between both intervals is at a $\Delta\text{IPD} = \pi$ radians but the highest proportions correct are usually at a lower ΔIPD. We interpret this as an indication of an increase in σ with IPD.

We aim to characterize the increasing internal processing uncertainty σ towards π radians, i.e. towards $\Delta\text{IPD} = 2\pi$ radians. This is done by fitting a periodic function to $\sigma(\text{IPD})$ and investigating the fit parameters. Based on Yost (1974), we constrained the function to have a minimum σ_0 (highest sensitivity) at IPD $= 0$ and a maximum at IPD $= \pi$ radians. A simple ad-hoc choice for such a function with one additional parameter is

$$\sigma(\text{IPD}) = g_\mathrm{m}\left[1 - \cos\left(\text{IPD}\right)\right] + \sigma_0, \tag{5.7}$$

with g_m as a measure of the magnitude of the IPD dependence. For better visualization, we excluded the frequency dependence for σ, σ_0 and g_m from Eq.(5.7). Figure 5.3 B shows $\sigma(\text{IPD})$ for three different values of g_m. Their influences on sensitivity are shown in Figure 5.3 C + D. We propose this descriptive model for the IPD sensitivity index:

$$\hat{d}'(\Delta\text{IPD}) = \frac{\Delta\mu(\Delta\text{IPD})}{\sigma(\Delta\text{IPD})} = \frac{\sin(\text{IPD})}{g_\mathrm{m}\left[1 - \cos(\text{IPD})\right] + \sigma_0}. \tag{5.8}$$

Again, a grid search was used to determine the parameters g_m (steps of 0.1 from 0 to 3) and σ_0 (steps of 0.1 from 0.1 to 0.8), which maximized the likelihood. The resulting functions for all stimulus conditions and subjects are displayed in Fig. 5.4 as solid lines.
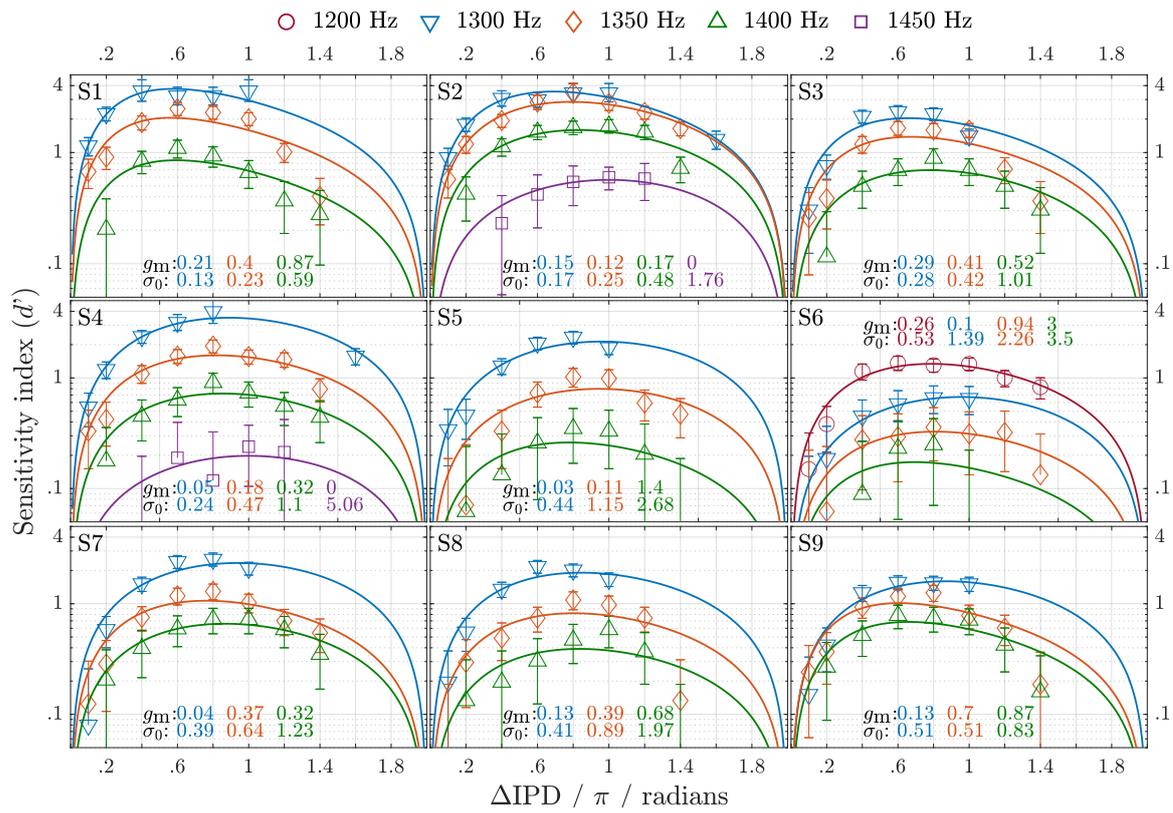
Figure 5.4.: Proportions correct from Fig. 5.2 transformed to d' (only for frequencies with more than 2 data points). Additionally, the fitted psychometric functions from Eq. (5.8) are plotted for each frequency and subject.

5.7. Analysis of the frequency dependence

With the data and the fitted psychometric functions, we were able to analyze IPD sensitivity as a function of frequency. Our foremost interest is the amount of decline in sensitivity, i.e. of $d'(f)$. In Fig. 5.5 the sensitivity index for different ΔIPDs is shown as a function of frequency. Those data were already shown as functions of ΔIPD in Fig. 5.4 and are now re-plotted as functions of f. We can directly use these d' values to estimate their decline across frequency. Alternatively, we can use the σ values, depending on σ_0 and g_{m} which were independently optimized (via maximum likelihood) for each frequency, from the previous fitted psychometric functions since $d' \sim 1/\sigma$. In the following, we will elaborate on the two ways to quantify the decline of d' with respect to frequency:

5.7.1. Likelihood fit on frequency dependence for specific IPD
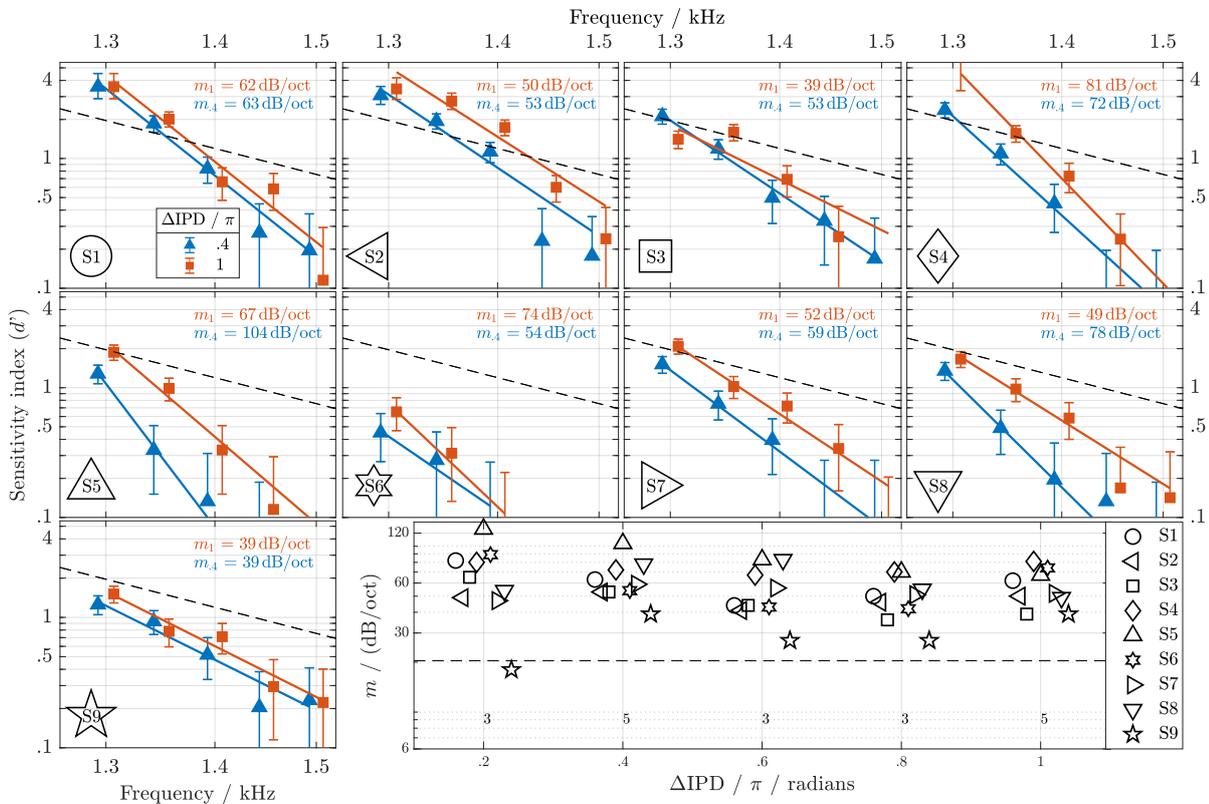
Figure 5.5.: ΔIPD-specific estimates of the frequency dependence of d'. **Panels S1-S9**: Symbols represent the d' values for different ΔIPDs across frequency of subjects S1 - S9. The error bars denote the 95 % confidence intervals. Solid lines show the maximum likelihood fit obtained from Eq. (5.6), see Section 5.7.1. d' prediction of the centroid model is depicted by the dashed black line (Brughera *et al.*, 2013). **Bottom right**: Decline m of $d'(f)$ for different ΔIPDs for the nine different subjects marked in symbols. The numbers below the symbols state data from how many frequencies were available for the fitting at the respective ΔIPD. The dashed black line represents the prediction of the centroid model.

The decline m in Eq. (5.6) is directly fitted to the sensitivity index as a function of frequency for a specific ΔIPD and subject. Again, a grid-search was used to maximize the likelihood. m ranges from 6 to 140 dB/oct in steps of 1 dB/oct. Since a higher decline requires a higher

intersection b, we set $b = mc$ where c ranges from 9.3 to 10.6 in steps of 0.0025. The solid lines in the 9 panels for the subjects of Fig. 5.5 represent the functions resulting in the maximum likelihood and the solutions for m are shown in the lowest right panel in Fig 5.5 for the nine subjects. Especially for the ΔIPDs with the most data points ($0.4\,\pi$ and $1\,\pi$), the declines are most consistent across subjects, resulting in 95% confidence intervals of 53-70 and 46-66 dB/oct, respectively. This shows the importance of the data points at 1450 and 1500 Hz in confining the decline estimate. Brughera *et al.* (2013) related a specific centroid measure of their model to a threshold to account for their data. Thus we were able to scale their measure to d' and calculate the decline across frequency (see Appendix 5.9.3). The d' decline over frequency from the centroid model (Brughera *et al.*, 2013) (dotted black line, 20 dB/oct) is too shallow to explain the sensitivity measured in this experiment.

5.7.2. Least-square fit on derived psychometric functions

As the decline of d' across frequency does not greatly depend on ΔIPD, the second approach is to derive a single decline value for each subject, combining data from different ΔIPD. Eq. (5.7) depicts the influence of the IPD on σ. To fit an IPD-independent decline we derived the mean value of σ over ΔIPD, which is

$$\overline{\sigma}(f) = \sigma_0(f) + g(f) = \sigma(f, \pi/2) \tag{5.9}$$

and is depicted in Fig 5.6 A. We used $d' = 1/\overline{\sigma}(f)$ with $f \geq 1300\,\text{Hz}$ (see example data in Fig. 5.6 B) to compute the least-square solution of m in Eq. (5.6), which is shown in Fig. 5.6 C for the nine subjects. The 95% confidence interval of the decline across subjects is 53-78 dB/oct.
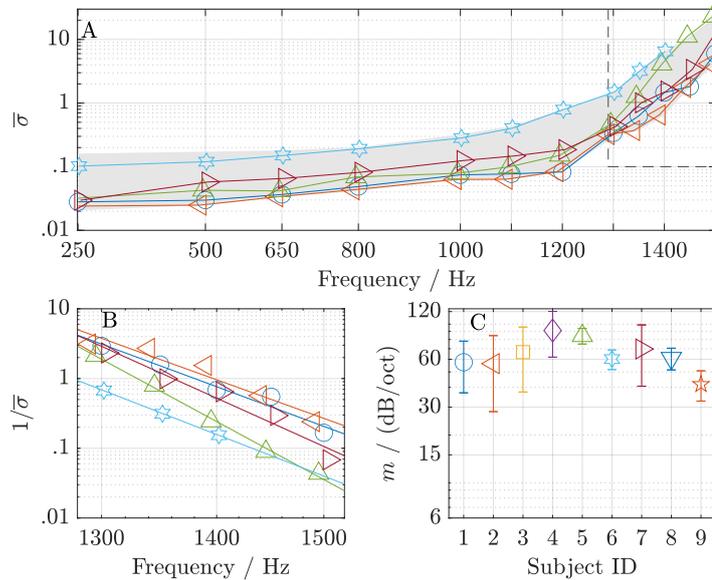
Figure 5.6.: Mean values of the fitted σ values and their decline with respect to frequency. **A** Range of $\overline{\sigma}$ for all subjects across frequency (gray area), and for some representative subjects (color code in C) selected to span the range of performance. **B** Symbols represent the same values as in A (color code in C), and solid lines represent the least-squares linear fit to the data for $f \geq 1300\,\text{Hz}$. **C** Decline of the linear fit in B. The error bars denote the 95 % confidence intervals.

5. Frequency dependence of sensitivity to interaural phase differences in pure tones

Both ways to calculate the decline across frequency clearly show that for all subjects it is greater than 30 dB/oct, see Fig. 5.6 C and Fig. 5.5 (bottom right) and the confidence intervals are mostly overlapping with an overall range of 46-78 dB/oct.

5.8. Discussion

5.8.1. ITD thresholds

To compare our behavioral results and the fitted psychometric functions to previous behavioral data, we calculated the ΔIPD thresholds ($p_c = 0.794 \Rightarrow d' = 1.16$) numerically from the best-fit psychometric functions and plotted the respective ΔITD thresholds in Fig. 5.7. Compared to previously published data also shown in Fig. 5.7, the thresholds reported in the present study at 250 Hz are lower for most subjects, but somewhat higher around 800 Hz. Consequently, while previously published thresholds decline with a frequency increase from 250 up to about 800 Hz, our data is better described by a broad plateau ranging from 500 to 800 Hz and sometimes even to 1100 Hz. While we do not have a good explanation for this deviation, Brughera *et al.* (2013) reported that small changes in their procedure, e.g. the choice of test frequencies, influenced the decline. Minor procedural differences that we consider unlikely to cause the difference are (1) the duration of 300 ms with 50 ms ramps compared to 500 ms duration with 100 ms ramps (Brughera *et al.*, 2013) and (2) that we used multiplicative step sizes in our adaptive procedure. The overall magnitude of thresholds and their steep increase above 1.2 kHz is in line with previously published data (see Fig. 5.7).
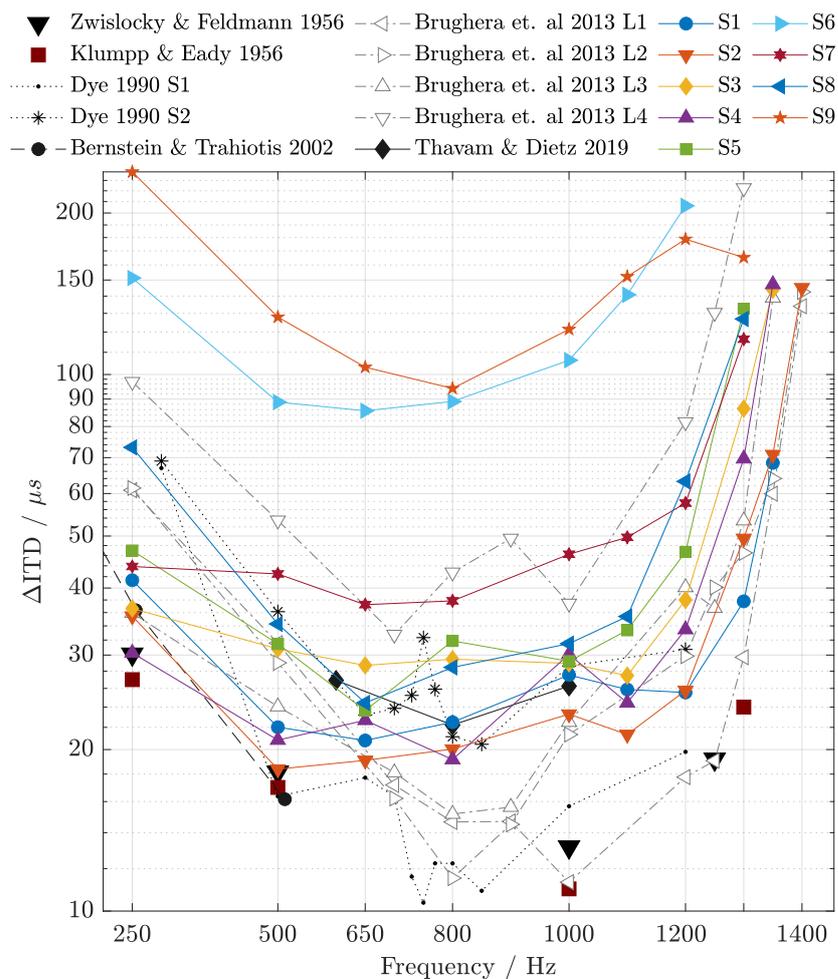
Figure 5.7.: Threshold ΔITDs corresponding to $d' = 1.16$ ($p_c = 0.794$) from subjects S1 through S9. Additionally, threshold ΔITDs are shown for previous data sets.

5.8.2. IPD dependence of sensitivity

There is a difference between the ΔIPD for the largest stimulus difference between both intervals, and where subjects gained the maximal correct proportion. Maximal sensitivity at ΔIPD $< \pi$ radians is in agreement with observations made by Yost (1974) who reported increasing thresholds with increasing reference IPD towards π radians. It also fits with the IPD dependence of Fisher information extracted from a population of ITD-sensitive neurons typical for mammals (Pavão *et al.*, 2020).

5.8.3. Steepness of IPD sensitivity decline

The results quantify the previously reported sudden reduction of $\mathrm{ITD}_{\mathrm{TFS}}$ sensitivity from $f = 1300\,\mathrm{Hz}$ to $f = 1500\,\mathrm{Hz}$ to be 9.5-16 dB. This corresponds to a d'-decline m of 46-78 dB/oct which is close to the decline derived from the behavioral data of Brughera *et al.* (2013) when assuming Eq. (5.5) (see Appendix 5.9.2). Such a steep decline is outstanding in psychoacoustics and is not assumed or reproduced by any model. In the following, we discuss mechanisms that may contribute to the decline, and are partially already implemented in existing models.

As described in the introduction, the reduction in sensitivity is commonly associated with the loss of synchrony of the AN to the TFS towards higher frequencies. Joris and Verschooten (2013) and Verschooten *et al.* (2019) directly connected the upper frequency limit of AN phase-locking with the upper frequency limit of $\mathrm{ITD}_{\mathrm{TFS}}$ sensitivity, and thus expect the limit of phase-locking in humans to be near 1400 Hz. Many phenomenological models of binaural processing include a low-pass filter to model the loss of synchrony of AN input (Bernstein and Trahiotis, 1996; Breebaart *et al.*, 2001a,b; Bouse *et al.*, 2019; Dietz *et al.*, 2009). The order of the filter is often taken from physiological studies. As the biophysical mechanisms causing the synchrony decline are arguably the same across species, the steepness is indeed very consistent across different species (Weiss and Rose, 1988a) and there is no reason to assume a different steepness in humans. The species-dependent corner frequency is then used as a free parameter to fit a model to human data (e.g., Bernstein and Trahiotis, 1996). Breebaart *et al.* (2001a,b) employed a cascade of five 1st-order low-pass filters with a resulting corner frequency at 770 Hz. In contrast to a single 5th-order low-pass filter, the cascaded design has a more gradual transition to maximum steepness and thus only 2 dB attenuation in the frequency interval 1300 to 1500 Hz. Therefore, their model still predicts low threshold ITDs at 1.5 kHz. This modest decline is in contrast to the steep $\mathrm{ITD}_{\mathrm{TFS}}$ sensitivity decline observed in the present study and to the data from Brughera *et al.* (2013). It can thus be concluded that additional effects must cause the steep decline of $\mathrm{ITD}_{\mathrm{TFS}}$ sensitivity. The next processing stage is the anteroventral cochlear nucleus (AVCN), which receives input from the AN and projects into the medial superior olive (MSO), where the first binaural interaction takes place. AVCN neurons preserve and further enhance the precision of temporal information in the neural firing of the AN (Pickles, 2015; Joris *et al.*, 1994). But the asymptotic synchrony decline across frequency in Joris *et al.* (1994) is even somewhat shallower than the average decline in AN fibers. Therefore, processing in the AVCN does not appear to contribute to the steeper roll-off of $\mathrm{ITD}_{\mathrm{TFS}}$ sensitivity.

Processing of the inputs within the MSO could be another origin of the decline of $\mathrm{ITD}_{\mathrm{TFS}}$ sensitivity across frequency. Even if its inputs were perfectly phase-locked, the duration of the input conductance imposes a frequency limitation on ITD sensitivity. The excitatory postsy-

naptic potential or the coincidence-detection window acts as an additional low-pass filter. This low-pass characteristic is also reflected in the modulation depth of the Brughera *et al.* (2013) model rate-ITD functions (already including the AVCN decline), declining by 3 dB between 1250 and 1500 Hz, i.e. a decline of about 2.4 dB can be expected in the 1300 to 1500 Hz interval. This decline is 1.4 dB more than the AVCN phase-locking decline of their model within this interval. Bouse *et al.* (2019) used the peripheral filter from Breebaart *et al.* (2001a) and include an additional low-pass filter within their MSO-model. Between 1300 and 1500 Hz, the ITD_{TFS} sensitivity of their model declines by 4 dB.

A common model approach for ITD encoding relies on an array of binaural coincidence-detecting neurons receiving differently delayed inputs from the left and right ear (Jeffress, 1948). Such a coincidence-detecting model unit responds maximally when the relative internal delay between its bilateral inputs exactly compensates for the external ITD. Based on this delay-line approach, binaural perception has commonly been modeled and explained by interaural cross-correlation (Jeffress, 1948; Colburn, 1977; Lindemann, 1986; Stern and Shear, 1996). To investigate this hypothesis, Brughera *et al.* (2013) examined the increase of the ITD_{TFS} thresholds across frequency with several model types, including a delay-line model (Jeffress, 1948) simulated as a delay-weighted (centroid) cross-correlation (Stern and Colburn, 1978). With increasing frequency, the number of cross-correlation cycles falling into the strongly-weighted delay range increases as well. As the weight of the first negative side-peak increases (with increasing ITD), it counteracts the primary peak, i.e. the one corresponding to the nominal ITD. This causes low-pass filtering and is one limitation of ITD_{TFS} sensitivity in this model. The sensitivity roll-off obtained with the model by Brughera *et al.* (2013) after centroid threshold calculation (thus including SI decline of AVCN and MSO low-pass filter) corresponds to a 4.2 dB attenuation for d' between 1300 and 1500 Hz. The p_{c} for the threshold IPD at 1400 Hz is 0.79 and would be 0.69 and 0.62 at 1500 and 1600 Hz, respectively. To our knowledge, this is the model with the steepest ITD_{TFS} sensitivity decline, but it is still far too gradual to account for the present data.

To summarize, the phase-locking of AN fibers declines relatively gradually (≤ 18 dB/oct), as does the simulated ITD sensitivity of auditory models (≤ 21 dB/oct). On the other hand, the ITD sensitivity index of our subjects declines in a narrow interval between 1300 and 1500 Hz by 46-78 dB/oct, unparalleled in psychoacoustics. This discrepancy shows that we do not yet fully understand the effects that lead to this steep decline. It is possible, that the upper limit of phase-locking in humans is at considerably higher frequencies (Moore, 2021; Verschooten *et al.*, 2019) and factors other than phase-locking are the primary cause of the frequency limitation of ITD sensitivity.

5.9. Appendix

5.9.1. Maximum likelihood fit

For the fitting in Sec. 5.5.2, 5.6.3 and 5.7.1 we used the following functions. To describe how well the observed k of N (proportion correct - p_c) fit to the probability p the binomial probability mass function serves as a likelihood function, where the likelihood is maximal if $p = p_c$:

$$l = b(k, N, p) = \binom{N}{k} p^k (1-p)^{N-k}, \tag{5.10}$$

with k the number of correct choices from N presentations. The likelihood function in Eq. (5.10) describes the 'goodness of fit' of the probability p for a single pair of k correct answers of N presentations. The likelihood for a series of this pairs is calculated by Eq. (5.11). Since these values could become small, and are not guaranteed to be numerically stable, the log-likelihood is used (5.12).

$$l = \prod_{j=1}^{J} b(k_j, N_j, \hat{p}_j), \tag{5.11}$$

$$\log(l) = \sum_{j=1}^{J} \log\left(b(k_j, N_j, \hat{p}_j)\right), \tag{5.12}$$

with j the index of data points and J the number of data points.

5.9.2. Sensitivity decline over frequency for a fixed IPD of the behavioral data by Brughera et al. (2013)

Brughera *et al.* (2013) reported threshold values for pure tones across frequency. These values show how ΔIPD changes for a fixed d' and are shown for their Listener 1 in Figure 5.8 A as a solid black line. However the goal in this study is to get the decline of how d' changes over a fixed ΔIPD. Therefore a description of the psychometric function is needed. Fitting Eq. (5.5) (with σ depending only on frequency) to the threshold values from Brughera *et al.* (2013) results in different sensitivity values (d') across frequency. The values of σ that provide the best fit to the thresholds at the three highest frequencies where thresholds were measurable in (Brughera *et al.*, 2013) for Listener 1, are shown in Fig. 5.8 B.

Obviously, the frequency range of considered data points influences the decline estimate. The focus of the present study is on the maximum decline that can be expected near the highest frequencies at which IPD sensitivity can be measured. The trade-off is that a too small interval and a fit at too high test frequencies increases the uncertainty of the decline estimate. A too large fit interval may result in an underestimation of the maximum decline. For example, the three data points from 1300 to 1400 Hz of their Listener 1 in Brughera *et al.* (2013) yield a decline of $m \approx 62$ dB/oct. Excluding the data-point at 1300 Hz yield a decline of $m \approx 66$ dB/oct, while including the data-point at 1250 Hz yield a decline of $m \approx 54$ dB/oct. Depending on frequency range and subject the sensitivity decline ranged from 30 dB/oct to 76 dB/oct.
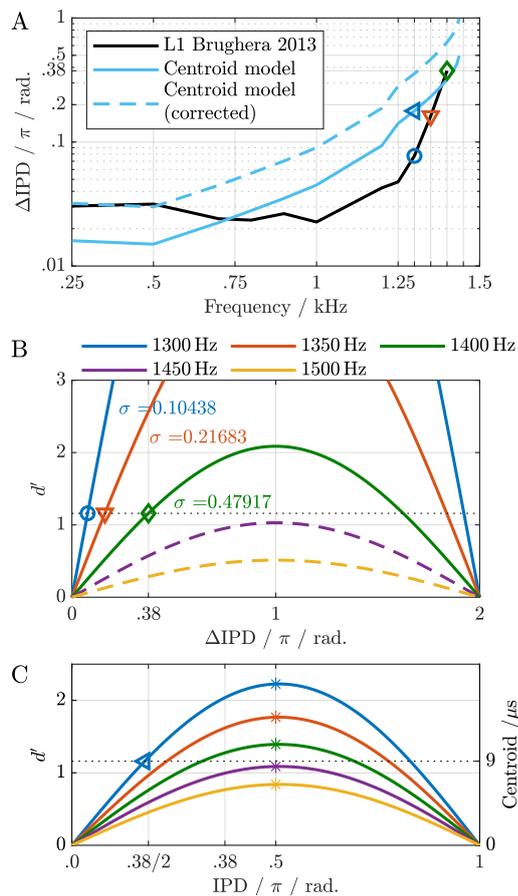
Figure 5.8.: **A**: Threshold ΔIPD corresponding to $d' = 1.16$ from Listener 1 (Brughera *et al.*, 2013) across frequency and predictions from the centroid model (Brughera *et al.*, 2013, Fig. 6a). The three markers represent the thresholds for 1300, 1350 and 1400 Hz. **B**: The psychometric function in Eq. (5.5) was matched to the threshold values from (A) by fitting σ for each frequency (solid lines). The vertical dotted line indicates $d'=1.16$. The dashed lines represent the psychometric function for 1450 and 1500 Hz assuming the decline of σ across frequency continues by -62 dB/oct. **C**: Output from centroid model in Brughera *et al.* (2013, Fig. 5). The horizontal dashed line (9 μs) marks the threshold criterion and thus relates to $d' = 1.16$. The intercepts with this line are equivalent with the threshold model predictions in (A). Despite the fact that the x-axis of this plot is IPD, Brughera *et al.* (2013) used this metric to derived ΔIPD threshold from the intercepts. This results in a factor of 2 error in the threshold estimates. The corrected (multiplication by two) predictions are plotted in (A) as dashed line.

5.9.3. Sensitivity decline over frequency for a fixed IPD of the centroid model by Brughera et al. (2013)

Brughera *et al.* (2013) used a delay-line model (Jeffress, 1948) simulated as a delay-weighted (centroid) cross-correlation (Stern and Colburn, 1978) to predict the ITD thresholds from their experiment. The rate ITD functions $c(\tau)$ of the coincidence-detecting model units are represented by the four parameter equation in Brughera *et al.* (2013, Appendix B). Values for the four parameters are reported for a set of frequencies. We did a linear interpolation for these values to obtain solutions for frequencies between 1300 to 1500 Hz. To calculate the centroid measure the density-weighted cross-correlation (Brughera *et al.*, 2013, Eq.(5)) was used:

$$\overline{\tau}(\Delta t) = \frac{\int p(\tau)\tau c(\tau - \Delta t)\ d\tau}{\int p(\tau)c(\tau - \Delta t)\ d\tau}, \tag{5.13}$$

with the density function $p(\tau)$ from Brughera *et al.* (2013, Eq.(6) & (7)). For the numerical solution we calculated the integrals as a sum with -2500 μs $\leq \tau \leq$ 2500 μs in steps of 1 μs.

Fig. 5.8 C shows the centroid across IPD for different frequencies (cf. Brughera *et al.*, 2013, Fig. 5). Brughera *et al.* (2013) relate the intersection of the centroid with 9 μs to the threshold value of $p_{\mathrm{c}} = .794$ (shown in Fig.5.8 A). This relation is equivalent to a linear scaling to d' (see left and right y-axis in Fig. 5.8 C). They extracted a ΔITD threshold from an ITD axis (here shown as IPD), so their predictions appear to be off by a factor of two.

The decline m discussed in this work describes how d' changes over frequency for a fixed IPD, this is equivalent to the change of d' in each vertical cut in 5.8 C. For an example IPD of $0.5\,\pi$ radians the d' values for the five frequencies in Fig. 5.8 C marked by stars are 2.23, 1.77, 1.39, 1.09 and 0.84. Solving Eq. (5.6) for these values results in a decline $m \approx 20\,\mathrm{dB/oct}$. This decline changes over IPD by a few tenths of a dB which is due to the numerical imprecision.

5.10. References

Bernstein, L. R., and Trahiotis, C. (**1996**). "The normalized correlation: Accounting for binaural detection across center frequency," J. Acoust. Soc. Am. **100**(6), 3774–3784, doi: `10.1121/1.417237`.

Bernstein, L. R., and Trahiotis, C. (**2002**). "Enhancing sensitivity to interaural delays at high frequencies by using "transposed stimuli"," J. Acoust. Soc. Am. **112**(3), 1026–1036, doi: `10.1121/1.1497620`.

Bernstein, L. R., and Trahiotis, C. (**2016**). "Behavioral manifestations of audiometrically-defined "slight" or "hidden" hearing loss revealed by measures of binaural detection," J. Acoust. Soc. Am. **140**(5), 3540–3548, doi: `10.1121/1.4966113`.

Bouse, J., Vencovský, V., Rund, F., and Marsalek, P. (**2019**). "Functional rate-code models of the auditory brainstem for predicting lateralization and discrimination data of human binaural perception," J. Acoust. Soc. Am. **145**(1), 1–15, doi: `10.1121/1.5084264`.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001a**). "Binaural processing model based on contralateral inhibition. i. model structure," J. Acoust. Soc. Am. **110**(2), 1074–1088, doi: `10.1121/1.1383297`.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001b**). "Binaural processing model based on contralateral inhibition. ii. dependence on spectral parameters," J. Acoust. Soc. Am. **110**(2), 1089–1104, doi: `10.1121/1.1383298`.

Brughera, A., Dunai, L., and Hartmann, W. M. (**2013**). "Human interaural time difference thresholds for sine tones: The high-frequency limit," J. Acoust. Soc. Am. **133**(5), 2839–2855, doi: `10.1121/1.4795778`.

Colburn, H. S. (**1977**). "Theory of binaural interaction based on auditory-nerve data. ii. detection of tones in noise," The Journal of the Acoustical Society of America **61**(2), 525–533, doi: `10.1121/1.381294`.

Dietz, M., Ewert, S. D., and Hohmann, V. (**2009**). "Lateralization of stimuli with independent fine-structure and envelope-based temporal disparities," J. Acoust. Soc. Am. **125**(3), 1622–1635, doi: `10.1121/1.3076045`.

Dye, R. H. (**1990**). "The combination of interaural information across frequencies: Lateralization on the basis of interaural delay," J. Acoust. Soc. Am. **88**(5), 2159–2170, doi: `10.1121/1.400113`.

Ewert, S. D. (**2013**). "AFC—A modular framework for running psychoacoustic experiments and computational perception models," in *Proceedings of the international conference on acoustics AIA-DAGA*, pp. 1326–1329.

Goldberg, J. M., and Brown, P. B. (**1969**). "Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization.," Journal of Neurophysiology **32**(4), 613–636, doi: `10.1152/jn.1969.32.4.613`.

Green, D. M., and Swets, J. A. (**1966**). *Signal Detection Theory and Psychophysics* (Wiley, New York).

Hafter, E. R., Dye, R. H., and Gilkey, R. H. (**1979**). "Lateralization of tonal signals which have neither onsets nor offsets," J. Acoust. Soc. Am. **65**(2), 471–477, doi: `10.1121/1.382346`.

Henning, G. (**1983**). "Lateralization of low-frequency transients," Hearing Research **9**(2), 153–172, doi: `10.1016/0378-5955(83)90025-4`.

Jeffress, L. (**1948**). "A place theory of sound localization," Journal of comparative and physiological psychology **41**(1), 35—39, doi: `10.1037/h0061495`.

Johnson, D. H. (**1980**). "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," The Journal of the Acoustical Society of America **68**(4), 1115–1122, doi: `10.1121/1.384982`.

Johnson, N. L., Kotz, S., and Kemp, A. W. (**1993**). *Univariate Discrete Distributions* (Wiley-Interscience, Hoboken, NJ).

Joris, P. X., Carney, L. H., Smith, P. H., and Yin, T. C. (**1994**). "Enhancement of neural synchronization in the anteroventral cochlear nucleus. i. responses to tones at the characteristic frequency," Journal of Neurophysiology **71**(3), 1022–1036, doi: `10.1152/jn.1994.71.3.1022`.

Joris, P. X., and Verschooten, E. (**2013**). "On the limit of neural phase locking to fine structure in humans," in *Basic Aspects of Hearing*, edited by B. C. J. Moore, R. D. Patterson, I. M. Winter, R. P. Carlyon, and H. E. Gockel, Springer New York, New York, NY, pp. 101–108, doi: `10.1007/978-1-4614-1590-9_12`.

Kiang, N. Y., Watanabe, T., Thomas, C., and Clark, L. (**1965**). *Discharge Patterns of single fibers in the cat's auditory nerve* (M.I.T. Press. Research Monograph 35, Cambridge, MA).

Klug, J., Encke, J., and Dietz, M. (**2023**). "Characterization of the decline in the auditory nerve phase locking at high frequencies," JASA Express Letters **3**(7), 074403, doi: `10.1121/10.0020267`.

Klumpp, R. G., and Eady, H. R. (**1956**). "Some measurements of interaural time difference thresholds," J. Acoust. Soc. Am. **28**(5), 859–860, doi: `10.1121/1.1908493`.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**(2B), 467–477, doi: `10.1121/1.1912375`.

Lindemann, W. (**1986**). "Extension of a binaural cross-correlation model by contralateral inhibition: I. Simulation of lateralization for stationary signals," The Journal of the Acoustical Society of America **80**(6), 1608–1622, doi: `10.1121/1.394325`.

Mills, A. W. (**1958**). "On the minimum audible angle," J. Acoust. Soc. Am. **30**(4), 237–246, doi: `10.1121/1.1909553`.

Moore, B. C. (**2021**). "Effects of hearing loss and age on the binaural processing of temporal envelope and temporal fine structure information," Hearing Research **402**, 107991, doi: `10.1016/j.heares.2020.107991` special Issue on Presbycusis.

Pavão, R., Sussman, E. S., Fischer, B. J., and Peña, J. L. (**2020**). "Natural ITD statistics predict human auditory spatial perception," Elife **9**, doi: `10.7554/eLife.51927`.

Pickles, J. (**2015**). "Auditory pathways: Anatomy and physiology," Handbook of clinical neurology **129C**, 3–25, doi: `10.1016/B978-0-444-62630-1.00001-9`.

Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (**1967**). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey.," Journal of Neurophysiology **30**(4), 769–793, doi: `10.1152/jn.1967.30.4.769`.

Stern, R. M., and Colburn, H. S. (**1978**). "Theory of binaural interaction based on auditory-nerve data. IV. A model for subjective lateral position," J. Acoust. Soc. Am. **64**(1), 127–140, doi: `10.1121/1.381978`.

Stern, R. M., and Shear, G. D. (**1996**). "Lateralization and detection of low-frequency binaural stimuli: Effects of distribution of internal delay," J. Acoust. Soc. Am. **100**(4), 2278–2288, doi: `10.1121/1.417937`.

Thavam, S., and Dietz, M. (**2019**). "Smallest perceivable interaural time differences," J. Acoust. Soc. Am. **145**(1), 458–468, doi: `10.1121/1.5087566`.

Verschooten, E., Robles, L., and Joris, P. X. (**2015**). "Assessment of the limits of neural phase-locking using mass potentials," Journal of Neuroscience **35**(5), 2255–2268, doi: `10.1523/JNEUROSCI.2979-14.2015`.

Verschooten, E., Shamma, S., Oxenham, A., Moore, B., Joris, P., Heinz, M., and Plack, C. (**2019**). "The upper frequency limit for the use of phase locking to code temporal fine structure in humans: A compilation of viewpoints," Hearing Research **377**, doi: `10.1016/j.heares.2019.03.011`.

Weiss, T., and Rose, C. (**1988**a). "A comparison of synchronization filters in different auditory receptor organs," Hearing Research **33**(2), 175–179, doi: `10.1016/0378-5955(88)90030-5`.

Weiss, T., and Rose, C. (**1988**b). "Stages of degradation of timing information in the cochlea: A comparison of hair-cell and nerve-fiber responses in the alligator lizard," Hearing Research **33**(2), 167–174, doi: `10.1016/0378-5955(88)90029-9`.

Yost, W. A. (**1974**). "Discriminations of interaural phase differences," J. Acoust. Soc. Am. **55**(6), 1299–1303, doi: `10.1121/1.1914701`.

Zwislocki, J., and Feldman, R. S. (**1956**). "Just noticeable differences in dichotic phase," J. Acoust. Soc. Am. **28**(5), 860–864, doi: `10.1121/1.1908495`.

6. Exploration of the abrupt decline in binaural pure tone sensitivity at high frequencies

In the previous chapter, we determined the decline of the IPD sensitivity for pure tones across frequency (46-78 dB/oct). In this chapter, we will discuss possible processes and mechanisms that cause this steep decline. An important mechanism is the decline of vector strength (VS) in the auditory nerve (AN) fibers towards higher frequencies ($\leq$18 dB/oct, Chapter 4). Another consideration is the head size and the ambiguities it creates in the ITD. The chapter focuses on two hypotheses: IPD sensitivity decreases due to peripheral filtering around 900 HZ (dominance region) and decreases due to synaptic low-pass filtering at the input of the MSO.

This chapter represents only a small fraction of the extensive research conducted over the past years. Numerous subdomains were explored and investigated by various researchers. Below are some specific contributions, some of them not presented in detail here:

Jonas Klug took the lead in organizing and driving the study, fostered interdisciplinary collaboration among various researchers, created graphics and charts and performed the calculations.

Jörg Encke's explanations have been instrumental in understanding the postsynaptic potential as a low-pass filter.

Go Ashida has contributed a lot of knowledge and practice in implementing and modeling the synaptic filter.

Helen Heinermann's detailed modeling of a spiking medial superior olive has contributed to the understanding of synaptic-filtering and its interaction with various model parameters.

Henri Pöntynen's diverse perspectives have always challenged and encouraged our conceptual outlook.

Bernhard Eurich has steadily contributed with his advice and broadened the perspective on psychoacoustic experiments with various interaural correlations.

Mathias Dietz played a pivotal role and provided significant support and assistance throughout the research process.

6.1. Introduction

The binaural temporal fine structure (TFS) sensitivity declines abruptly above 1400 Hz (Chapter 5, Brughera *et al.*, 2013; Klug and Dietz, 2022). Klug and Dietz (2022) indicate a range of 46-78 dB/oct (153-260 dB/dec) for the subjects studied. This decline, which we often refer to as "cliff-like", is very steep and unparalleled in psychoacoustics. Figure 6.1 shows the idealized range of proportion correct in a behavioral IPD discrimination experiment across frequency for a ΔIPD of $0.8\,\pi$ radians. We are interested in understanding which processes contribute to this steep decline across frequency. In the following, four present hypotheses on the origin of the decline are discussed.
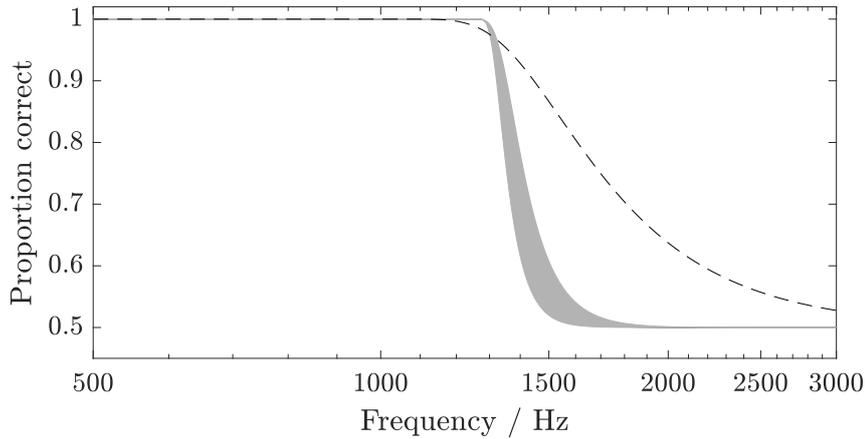
Figure 6.1.: Average range of the proportion correct across frequency for a pure tone with a ΔIPD of $0.8\,\pi$ radians (gray area) declining with 46-78 dB/oct. Proportion correct refers to the detected direction of movement when the IPD changes between two intervals (cf. Section 5.4.2, Klug and Dietz, 2022). The dashed line sketches a hypothetical psychometric function based purely on the AN decline (12 dB/oct).

The periphery hypothesis

A common hypothesis on the mechanism behind the decrease in sensitivity is that mainly the decrease in phase locking of the auditory nerve (AN) fibers is the origin (e.g. Verschooten *et al.*, 2019). However, the analyses in Chapter 4 (Klug *et al.*, 2023) shows that the loss of phase locking in the AN is only $\leq$18 dB/oct (or $\leq$60 dB/dec). This decline is not sufficient to account for the steep decline of ITD$_{\text{TFS}}$ sensitivity. A psychometric function that would be based on the AN decline is sketched as a dashed line in Figure 6.1.

The head size argument

Another possible explanation is that later stages have adapted to suppress the information because the spatial information carried by ITD_{TFS} becomes ambiguous above a certain frequency (depending on head size). The high-precision neurons that are important for processing TFS sensory information require a significant amount of energy to function properly. Given the high metabolic cost of encoding performed by these cells, it seems unlikely that the information they carry at high frequencies is simply discarded. Furthermore the spatial information extracted from ITD_{TFS} becomes spatially ambiguous at much lower frequencies than 1400 Hz, and the information provided by these neurons can still be useful when combined with other sources of information, such as ILD.

Ultimately, this hypothesis could only be tested by comparing the upper frequency limit at different head sizes. This would require species with different head sizes, as the differences within a species are likely to be too small. An alternative would be to test the limit on artificial neural networks trained for localization with different ear distances. However, it is doubtful whether these networks would reproduce a steep limit.

The synaptic-filter hypothesis

From our modeling work we concluded that the duration and shape of the excitatory postsynaptic potential (EPSP) of the MSO (also called synaptic filter or coincidence-detection window) acts like a low-pass filter and contributes to a steeper decline of IPD sensitivity (Jörg Encke, Ashida *et al.*, 2013; Heinermann *et al.*, 2019). In addition to the factors outlined above, the widely discussed inhibitory input to the MSO (Cant and Hyson, 1992; Grothe and Sanes, 1993; Brand *et al.*, 2002; Pecka *et al.*, 2008; Jercog *et al.*, 2010) hypothetically contributes to the steepness of the decline. In this chapter we model the decline introduced by the synaptic filter.

6. Exploration of the abrupt decline in binaural pure tone sensitivity at high frequencies

The dominance-region hypothesis

Some models (e.g. Brughera *et al.*, 2013) account for the abrupt sensitivity decline with an explicit change of parameterization between 1400 and 1500 Hz. It could also be argued that the binaural interaction that encodes the TFS only occurs up to a certain frequency. Folkerts and Stecker (2022) discovered an ITD *dominant region* between 600-800 Hz where ITD is most valuable for lateralization and localization of broadband stimuli. Based on this finding Goupell *et al.* (2023) hypothesised that the steep decline in ITD sensitivity reflects the peripheral frequency tuning of the neurons within the dominant region. In all mentioned cases, high-frequency signals would be processed by lower channels where normal binaural processing takes place and the sensitivity should decrease with the slope of the auditory band-pass filter. In fact, peripheral filter functions (e.g. Hohmann, 2002; Zwicker, 1974) decrease in the same order of magnitude as the decrease in binaural sensitivity. Figure 6.2 shows two band-pass peripheral filters from Hohmann (2002) with center frequency of 970 Hz (blue) and 1300 Hz (red): both decrease towards higher frequencies and match the decline range of the behavioral IPD sensitivity (Klug and Dietz, 2022) marked by the gray area. Yin and Chan (1990) presented Rate-IPD functions for a single MSO cell for different frequencies that cover the response area of that cell (shown in Figure 2.7 A). This cells rate magnitude is proportional to IPD sensitivity. The implicit bandpass filter of this magnitude function is compatible with the IPD sensitivity decline (dominant region).

Following the dominance-region hypothesis, the sensitivity around the upper limit should be strongly level-dependent and for high levels a significant sensitivity for frequencies $\geq$1500 Hz should be detectable. While ITD sensitivity seems to be independent of stimulation level for broadband stimuli (Thavam and Dietz, 2019), this is not the case for sinusoidal signals. Instead, performance in pure tone ITD discrimination tasks increases significantly with increasing stimulation level both in the area of the upper limit (Goupell *et al.*, 2023) but also for lower frequencies (Zwislocki and Feldman, 1956). It is important to emphasize that the level effect is not exclusive to the high frequencies. It remains unclear whether the magnitude of the level dependence varies with stimulation frequency, and what implications this variability would have on the current models of binaural hearing.

The finding of Pecka *et al.* (2008), reporting best ITD sensitivity for frequencies $\leq$ 1200 Hz, even if the cells which they recorded from in gerbils had characteristic frequencies up to 4800 Hz, contradicts the hypothesis of a *dominant region*. In Section 6.2, an experiment that further disproves the dominance-region hypothesis is presented where the IPD sensitivity to a 1.3-kHz tone does not change in the presence of a binaurally uncorrelated, 1.1-kHz low-pass masking noise.
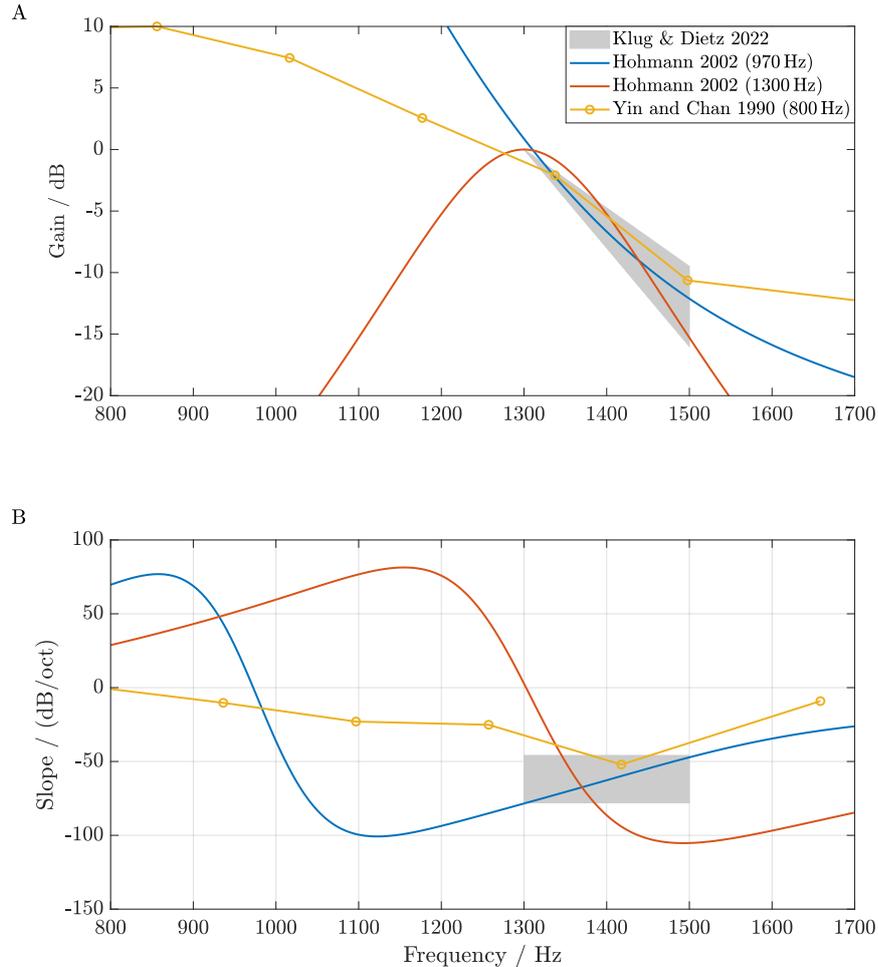
Figure 6.2.: Different band-pass filter functions matching the slope range (gray area) of IPD$_{\text{TFS}}$ sensitivity measured in Klug and Dietz (2022). **A** Gain across frequency. **B** Slope across frequency. The blue and red line show a filter with a different center frequency by Hohmann (2002) where the yellow line is derived from the magnitude of the Rate-IPD function of an MSO cell (Yin and Chan, 1990) across frequency (shown in Figure 2.7 B).

6.2. Experiment assessing the dominance-region hypothesis

According to the dominance-region hypothesis, the presence of binaural masking noise whose upper frequency limit is below the frequency of a high-frequency target tone (e.g. 1300 Hz) should eliminate any useful signal in the dominant region, while leaving the signal-to-noise ratio at the frequency of the target tone relatively unaffected. Consequently, if binaural information is derived mainly from the dominant region, subjects should either fail to perform the task or perform significantly worse than under conditions where the masking noise is absent.

In order to evaluate the dominance-region hypothesis the threshold IPD of a 1300 Hz pure tone was measured in the presence and absence of binaurally uncorrelated masking noise (Figure 6.3 A). In the masker-noise condition, the tone was presented along with a lowpass noise up to 1100 Hz.

6.2.1. Procedure

The procedure was a two-interval, two-alternative forced-choice task (2I-2AFC). The 300-ms tone duration included 50-ms $\cos^2$ rise-decay ramps. A 50-ms silent interval separated the two intervals. A next pair of intervals was presented 500 ms after the subject responded. Visual feedback was provided after each trial. The subject was required to respond according to whether the tone in the second interval was perceived to the left or the right of the tone presented in the first interval. The stimuli had synchronous onset and offset gating in both ears, so that the tones differed only in their IPD.

Tones were presented to the subject via Sennheiser HD-650 headphones at a level of 65 dB SPL. The subjects were seated in a double-walled, sound-attenuating booth and responded by pressing a key on a standard computer keyboard. A total of five self reported normal-hearing subjects aged between 21 and 35 years (F = 3, M = 2) participated in the experiment.

The interaural difference of the stimuli presented in the two intervals were symmetrical around zero as in many previous experiments on ITD sensitivity (e.g. Hafter *et al.*, 1979; Henning, 1983; Brughera *et al.*, 2013; Thavam and Dietz, 2019), so that in one of the two intervals, the right ear was leading, while in the other interval the left ear led by the same IPD. Subjects could thus make their decision based on a ΔIPD difference between the two intervals.

An adaptive 'three-down, one-up' staircase procedure controlled the ΔIPD, i.e. the ΔIPD was decreased after three correct responses in a row and increased after each incorrect response, asymptoting a p_c of 0.707 (Levitt, 1971). Each adaptive track started at a ΔIPD of $0.4\,\pi$ radians. The initial step size was a factor of 2, which was reduced to 1.414 and 1.189 after the first and second 'down-up-reversal'. An adaptive track was terminated after 10 reversals at the minimum step size.

6.2.2. Results
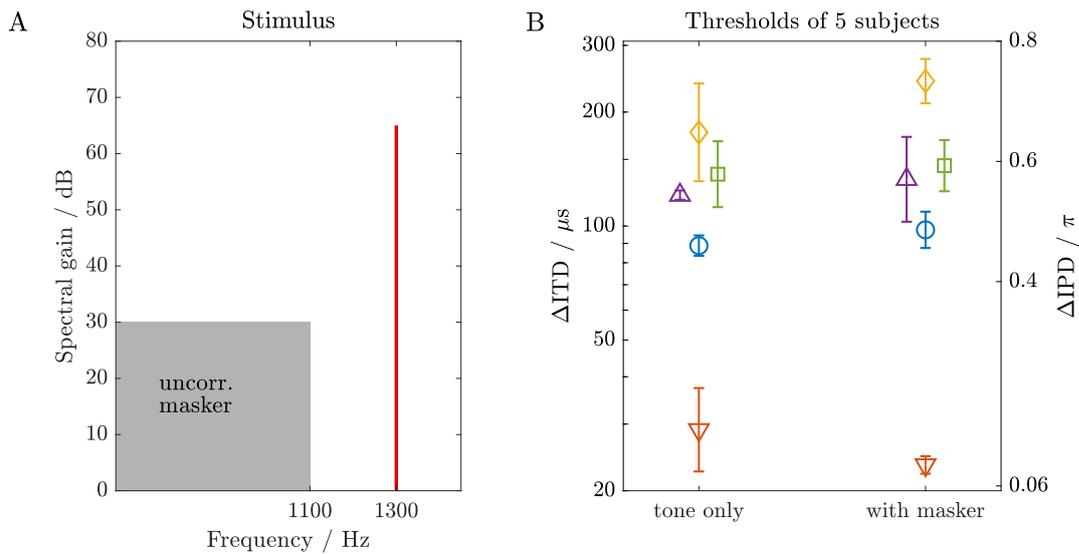
Figure 6.3.: **A** Spectrum of the Stimulus; **B** Thresholds of ΔITD sensitivity for the five subjects for tone only and tone with masker. The error bars denote the standard deviation.

Figure 6.3 B shows the ΔITD thresholds of all subjects. There is no systematic difference in

thresholds between the *tone only* and the *with masker* condition.

6.3. Modeling synaptic filtering

6.3.1. Modeling methods

For modelling EPSPs the alpha function (see Fig. 6.4 A)

$$\alpha(t) = (Ht/\tau) \exp(1 - t/\tau), \tag{6.1}$$

with H the magnitude and a half-peak width of 2.445τ, is widely used as a filter kernel that is convolved with an input (Silberberg *et al.*, 2004; Sterratt *et al.*, 2011; Ashida *et al.*, 2017). While the corner frequency of this synaptic filter depends on τ (Ashida *et al.*, 2013) the decline across frequency is about $11\,\mathrm{dB/oct}$ ($36\,\mathrm{dB/dec}$), see Fig. 6.4 B. Even when combined with the AN phase locking decline the resulting slope is not steep enough to account for the strong decline of pure tone IPD sensitivity.
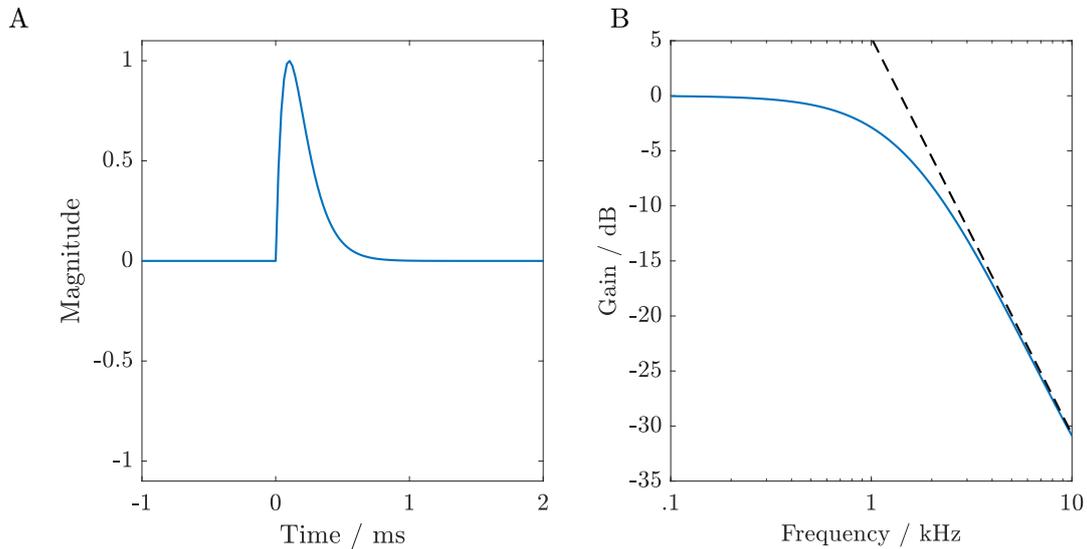
Figure 6.4.: **A** α-function from Eq. (6.1) as EPSP; **B** Transfer function of the EPSP. The dashed black line shows a decay of $-36\,\mathrm{dB/dec}$

Previous studies (e.g., Cant and Hyson, 1992; Grothe and Sanes, 1993) found that there are also inhibitory inputs to the MSO (as sketched in Figure 2.5), which introduce inhibitory postsynaptic potentials (IPSPs). For modeling this IPSP the alpha function from Eq. (6.1) is used, where H_i is a negative number to account for the inhibitory effect. The overall convolution kernel $k_c(t)$ is the sum of the EPSP and the IPSP:

$$k_c(t) = \alpha_e(t) + \alpha_i(t + \Delta t_i), \tag{6.2}$$

where Δt_i is the timing offset between EPSP and IPSP. Since the inhibitory input precedes the excitatory input (Brand *et al.*, 2002) $\Delta t_i > 0$. This additional IPSP shifts the peak of the IPD tuning functions (Brand *et al.*, 2002; Pecka *et al.*, 2008; Jercog *et al.*, 2010). The influence of its timing (Δt_i) on the IPD sensitivity is discussed in Roberts *et al.* (2013) and Myoga *et al.* (2014). Regardless of this, the interaction of an EPSP and an IPSP (or the resulting synaptic

filter as convolution kernel) also influences the decline across frequency. In the following model we investigated whether there is a plausible parameterization of the kernel that achieves the steepness of the IPD sensitivity observed by Klug and Dietz (2022). The constraint for this fit was the same energy (equivalent areas under the time domain kernel plot) for the EPSP and IPSP. Thus the magnitude of the IPSP was computed as

$$H_i = -\frac{\tau_e}{\tau_i}. \tag{6.3}$$

6.3.2. Modeling results

Although the overall decline does not change by adding the IPSP, the decline in certain frequency sections varies depending on the timing, half-peak width and amplitude of the IPSP. Because the frequency range within which the IPD sensitivity changes is narrow, the ripple in the frequency response of the synaptic filter can be adjusted to allow a drop of about $9\,\mathrm{dB}$ between 1300 and $1500\,\mathrm{Hz}$ (see red solid line in Figure 6.5). With $\tau_e = 300\,\mu s$, $\tau_i = 1.26\,\tau_e$ and a Δt_i of $650\,\mu s$ the resulting kernel has a slope of -45 dB/oct in the range between 1300 and $1500\,\mathrm{Hz}$. The VS decline across frequency at the AN contributes to the steepness of this decline and dampens the ripple at higher frequencies (see red dashed line in Figure 6.5).
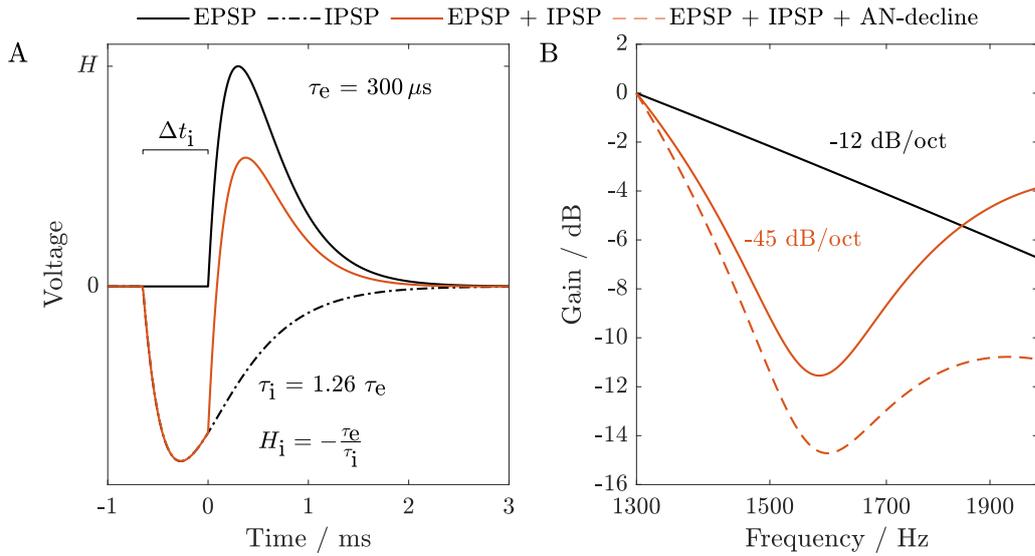
Figure 6.5.: **A** Kernel of the synaptic filter in the time domain; **B** Transfer function of the synaptic filter.

6.4. Discussion

With this conceptual proof, the synaptic filter at the input of the MSO (first binaural interaction) is a possible processing stage contributing to the steep decline in IPD sensitivity. Nevertheless, based on all considerations in this chapter, it seems reasonable to assume that the decline of behavioral IPD sensitivity is a result of the compounded low-pass filtering arising from the combined contributions of the individual stages of the auditory pathway. All hypothesis listed in Section 6.1 may contribute with a specific amount to the decline. According to Klug *et al.* (2023) we know that the steep decline of 46-78 dB/oct is not at the output of the AN. In Section 6.2, we showed that the dominance region is not pronounced enough to fully explain the decline. It remains unclear whether the strong decline is reached at the output of the binaural interaction (as suggested by Phil Joris) or only after further processing (cf. head-size hypothesis). An electrophysiological study at the MSO output in the critical frequency region would be necessary to further characterize the decline.

However, a steep low pass filtering by the compounded stages would have serious consequences on the neural signals delivered to subsequent processing stages. For instance, in the case of a pure tone stimulus, all of the higher harmonics introduced by the peripheral processing (see Figure 2.6 A) would be attenuated, resulting in a pure sine wave. With a sinusoidal rate-IPD function of the left and right MSOs the best sensitivity would occur at an ΔIPD of π. However, Chapter 5 (Klug and Dietz, 2022) clearly showed that the best sensitivity is at about 0.8 π for those frequencies in the steep decline range.

To allow for such behavior, the following scenarios are possible: (1) the harmonics are still present and there is no steep filter but trained decoding in the cortex, (2) nonlinearities in the decoding of sinusoidal rate-IPD functions, or (3) an IPD-dependent noise source that makes the decoding more uncertain for larger IPDs as suggested in Klug and Dietz (2022). For this trade-off between steepness and inclusion of higher harmonics, the dominance-region hypothesis seems to be a good solution because the peripheral filtering takes place before half-wave rectification, preserving the higher harmonics. However, ITD sensitivity in neurons with high CFs (Pecka *et al.*, 2008) and the results from this experiment provide some evidence against this hypothesis. Nevertheless, no clear conclusion can be drawn yet. A broader formal study based on this small experiment and also an extension of the methods (in accordance with the 'method triad', Figure 1.1) is needed, e.g., modeling AN fiber VS in the dominant region in the presence of noise.

6.5. References

Ashida, G., Funabiki, K., and Carr, C. (**2013**). "Biophysical basis of the sound analog membrane potential that underlies coincidence detection in the barn owl," Frontiers in Computational Neuroscience **7**, doi: `10.3389/fncom.2013.00102`.

Ashida, G., Tollin, D. J., and Kretzberg, J. (**2017**). "Physiological models of the lateral superior olive," PLOS Computational Biology **13**(12), 1–50, doi: `10.1371/journal.pcbi.1005903`.

Brand, A., Behrend, O., Marquardt, T., McAlpine, D., and Grothe, B. (**2002**). "Precise inhibition is essential for microsecond interaural time difference coding," Nature **417**(6888), 543–547, doi: `10.1038/417543a`.

Brughera, A., Dunai, L., and Hartmann, W. M. (**2013**). "Human interaural time difference thresholds for sine tones: The high-frequency limit," J. Acoust. Soc. Am. **133**(5), 2839–2855, doi: `10.1121/1.4795778`.

Cant, N. B., and Hyson, R. L. (**1992**). "Projections from the lateral nucleus of the trapezoid body to the medial superior olivary nucleus in the gerbil," Hearing research **58**(1), 26–34, doi: `10.1016/0378-5955(92)90005-8`.

Folkerts, M. L., and Stecker, G. C. (**2022**). "Spectral weighting functions for lateralization and localization of complex sounda)," The Journal of the Acoustical Society of America **151**(5), 3409–3425, doi: `10.1121/10.0011469`.

Goupell, M. J., Stecker, G. C., and Tollin, D. J. (**2023**). "Is the rapid decline in interaural time difference sensitivity above 700 Hz explained by downward spread of excitation into the frequency dominant region?," The Journal of the Acoustical Society of America **153**(3_supplement), A335–A335, `https://www.researchgate.net/publication/370307645_Is_the_rapid_decline_in_interaural_time_difference_sensitivity_above_700_Hz_explained_by_downward_spread_of_excitation_into_the_frequency_dominant_region`, doi: `10.1121/10.0019056`.

Grothe, B., and Sanes, D. H. (**1993**). "Bilateral inhibition by glycinergic afferents in the medial superior olive," Journal of Neurophysiology **69**(4), 1192–1196, doi: `10.1152/jn.1993.69.4.1192` pMID: 8492158.

Hafter, E. R., Dye, R. H., and Gilkey, R. H. (**1979**). "Lateralization of tonal signals which have neither onsets nor offsets," J. Acoust. Soc. Am. **65**(2), 471–477, doi: `10.1121/1.382346`.

Heinermann, H. T., Klug, J., Herrmann, S., Ashida, G., Encke, J., and Dietz, M. (**2019**). "On the frequency limit of interaural time difference sensitivity for pure tones," in *23. International Congress on Acoustics*, Aachen, Germany, doi: `10.18154/RWTH-CONV-239106`.

Henning, G. (**1983**). "Lateralization of low-frequency transients," Hearing Research **9**(2), 153–172, doi: `10.1016/0378-5955(83)90025-4`.

Hohmann, V. (**2002**). "Frequency analysis and synthesis using a gammatone filterbank," Acta Acustica united with Acustica **88**(3), 433–442.

Jercog, P. E., Svirskis, G., Kotak, V. C., Sanes, D. H., and Rinzel, J. (**2010**). "Asymmetric excitatory synaptic dynamics underlie interaural time difference processing in the auditory system," PLOS Biology **8**(6), 1–9, doi: `10.1371/journal.pbio.1000406`.

Klug, J., and Dietz, M. (**2022**). "Frequency dependence of sensitivity to interaural phase differences in pure tones," The Journal of the Acoustical Society of America **152**(6), 3130–3141, doi: `10.1121/10.0015246`.

Klug, J., Encke, J., and Dietz, M. (**2023**). "Characterization of the decline in the auditory nerve phase locking at high frequencies," JASA Express Letters **3**(7), 074403, doi: `10.1121/10.0020267`.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**(2B), 467–477, doi: `10.1121/1.1912375`.

Myoga, M. H., Lehnert, S., Leibold, C., Felmy, F., and Grothe, B. (**2014**). "Glycinergic inhibition tunes coincidence detection in the auditory brainstem," Nature Communications **5**, doi: `10.1038/ncomms4790`.

Pecka, M., Brand, A., Behrend, O., and Grothe, B. (**2008**). "Interaural time difference processing in the mammalian medial superior olive: The role of glycinergic inhibition," Journal of Neuroscience **28**(27), 6914–6925, doi: `10.1523/JNEUROSCI.1660-08.2008`.

Roberts, M. T., Seeman, S. C., and Golding, N. L. (**2013**). "A mechanistic understanding of the role of feedforward inhibition in the mammalian sound localization circuitry," Neuron **78**(5), 923–935, doi: `10.1016/j.neuron.2013.04.022`.

Silberberg, G., Wu, C., and Markram, H. (**2004**). "Synaptic dynamics control the timing of neuronal excitation in the activated neocortical microcircuit," The Journal of Physiology **556**(1), 19–27, doi: `10.1113/jphysiol.2004.060962`.

Sterratt, D., Graham, B., Gillies, A., and Willshaw, D. (**2011**). *Principles of Computational Modelling in Neuroscience* (Cambridge University Press).

Thavam, S., and Dietz, M. (**2019**). "Smallest perceivable interaural time differences," J. Acoust. Soc. Am. **145**(1), 458–468, doi: `10.1121/1.5087566`.

Verschooten, E., Shamma, S., Oxenham, A., Moore, B., Joris, P., Heinz, M., and Plack, C. (**2019**). "The upper frequency limit for the use of phase locking to code temporal fine structure in humans: A compilation of viewpoints," Hearing Research **377**, doi: `10.1016/j.heares.2019.03.011`.

Yin, T. C., and Chan, J. C. (**1990**). "Interaural time sensitivity in medial superior olive of cat," Journal of Neurophysiology **64**(2), 465–488, doi: `10.1152/jn.1990.64.2.465`.

Zwicker, E. (**1974**). "On a psychoacoustical equivalent of tuning curves," in *Facts and Models in Hearing*, edited by E. Zwicker and E. Terhardt, (Communication and cybernetics), Springer Verlag, pp. 132–141.

Zwislocki, J., and Feldman, R. S. (**1956**). "Just noticeable differences in dichotic phase," J. Acoust. Soc. Am. **28**(5), 860–864, doi: `10.1121/1.1908495`.

7. General conclusion

In this thesis, the processing of interaural differences in the mammalian brain was investigated with psychophysical experiments and computational modeling. Although the opponent-channel model is in many ways better motivated by mammalian physiology, historically, delay-line models have been more successful at accounting for a wide variety of binaural behavioral data. An opponent-channel model has been developed that is capable of predicting a wide range of human behavioral data (lateralization of ILDs and ITDs$_{ENV}$) without the use of a delay-line approach (Chapter 3). Only one pair of spiking model EI-neurons (composed of one neuron from the left and one from the right) was employed for each center frequency. This brings the opponent-channel model up-to-par with the delay-line models, at least in the sense of qualitative prediction of ILD and ITD$_{ENV}$ data.

This is an important building block in understanding binaural processing of envelope (ENV) information. It is of particular interest for cochlear implant (CI) users since they only have access to this ENV information due to the usage of coding strategies that only encode the same (cf. Williges *et al.*, 2018). In the modeling study of Hu *et al.* (2022), we employed a model of an electric driven auditory periphery (Hamacher, 2004; Fredelake and Hohmann, 2012) to our model chain: We were able to produce electrically driven rate-ITD functions and depict the influence of the pulse rate in the CI processing. This approach was extended to multi-channel electrodes and consideration of the whole tonotopy to investigate CI coding strategies (Hu *et al.*, 2023).

Furthermore, Chapter 3 (Klug *et al.*, 2020) showed that at high stimulation levels, commonly used in psychoacoustic experiments, off-frequency channels are essential for encoding ENV information. In the on-frequency auditory nerve (AN) fibers the vector strength (VS) to the ENV decreases with increasing level (Joris and Yin, 1992), while sensitivity to ITD$_{ENV}$ increases with level (Dietz *et al.*, 2013). One clear explanation for this apparent contradictions is the usage of information in off-frequency channels (Joris and Yin, 1992; Klug *et al.*, 2020; Johannesen *et al.*, 2022). The use of off-frequency information for binaural tasks was already discussed by Bernstein and Trahiotis (2008) and Dreyer and Oxenham (2008). Bernstein and Trahiotis (2008) could show that binaural discrimination was possible in the presence of spectrally flanking notched noise, which masks the off-frequency envelope information. This suggests that most of the ENV information is encoded in the on-frequency channel, contradicting the improving ITD$_{ENV}$ sensitivity. Our current ongoing work (Klug *et al.*, 2024) deals with the hypothesis that this off-frequency noise enhances the synchronization of the neural responses to the envelope in the on-frequency channel by influencing the outer hair cell (OHC) compression. Subjects would exploit off-frequency channels in the absence of notched noise but on-frequency channels in its presence. Using off-frequency channels has implications for hearing aid fitting. Compensating for hearing loss by amplifying in a particular frequency band affects the channels that carry the envelope information from that frequency band.

7. General conclusion

The biggest challenge in predicting ITD_TFS discrimination data for both the delay-line model or the opponent-channel model, is being able to convincingly relate the steep decline in IPD sensitivity at frequencies above 1.3 kHz to physiologically plausible sources. Previous research has shown that vector strength (VS) of AN fibers has a low-pass characteristic across frequency (Weiss and Rose, 1988; Verschooten *et al.*, 2015). This decline in the peripheral processing was often seen as the explanation for the decreasing ITD_TFS sensitivity (Joris and Verschooten, 2013; Verschooten *et al.*, 2019). In Chapter 4 (Klug *et al.*, 2023), the synchrony of AN fiber responses to the TFS was re-analyzed. We found that the slope of the VS decline across frequency is not as steep as suggested by the nominal filter orders used by the low-pass filters employed in existing models of the auditory periphery. A behavioral experiment (Chapter 5) has shown that the ITD_TFS sensitivity decline is much steeper than predicted by existing models. Therefore, we find it unlikely that the very steep decline observed in behavioral experiments of ITD sensitivity could be adequately explained solely by the low-pass behavior of VS across frequency. Possible reasons or mechanisms for this decline have been mentioned and discussed further in Chapter 6.

The most important results are listed in Section 7.1, while Section 7.2 provides insight into potential avenues for further exploration.

7.1. Summary & Conclusions

- We can explain extensive data sets of human ILD and ITD_ENV based lateralization with the two-channel concept (Klug *et al.*, 2020, Chapter 3).

- Integration of off-frequency channels is essential to reconcile physiological and psychoacoustic phenomena. (Klug *et al.*, 2020, Chapter 3).

- With a slope of approximately $12\,\text{dB/oct}$ ($40\,\text{dB/dec}$) the decline of AN phase locking across frequency is less steep than previously expected (Klug *et al.*, 2023, Chapter 4).

- We characterized the long-known decline in sensitivity to IPD_TFS across frequency as 46-78 dB/oct (153-260 dB/dec) and found it to be steeper than previously thought (Klug and Dietz, 2022, Chapter 5).

- Best IPD sensitivity for pure tones in the frequency range of 1300 to 1500 Hz was found at $\text{IPD}_\text{TFS} < \pi/2$ radians (Klug and Dietz, 2022, Chapter 5).

- We speculated on processing stages which contribute to the steep IPD sensitivity decline. For the synaptic filter we could show a partial decline of $45\,\text{dB/oct}$ (Chapter 6).

- Any strong low-pass filtering after the periphery removes higher harmonics from the signal and this favors best sensitivity at $\text{IPD}_\text{TFS} = \pi/2$, contradicting the previous finding (Chapter 6).

7.2. Suggestions for future research

The simulations in Chapter 3 (Klug *et al.*, 2020) reproduce human behavioral data and physiological rate ITD functions of LSO cells. Both the model (Figure 3.3(A)) and physiologic recordings (Figure 2.8 (B)) show that ILDs and $ITDs_{ENV}$ are encoded in the same rate space. This means that it is not possible to distinguish whether a certain rate is caused by an ILD or ITD_{ENV}. This is relevant to the question of whether the binaural cues are processed in independent channels. Furukawa (2008) has investigated the degree of binaural channel interaction by comparing behavioral sensitivity for combined cues with those for only ILD or ITD_{ENV}. For a model neuron in Klug *et al.* (2020), an interaction of a positive ITD_{ENV} and a negative ILD could cause the same rate as a stimulus without any binaural cues. Knowing whether there is such a stimulus constellation that humans cannot discriminate would be helpful in answering the above question. Of course, the CF must also be considered, as the model neurons at different CFs respond differently to ITD and ILD.

In Chapter 5, we characterized the decline of ITD_{TFS} sensitivity across frequency in detail. This characterization is an important aspect of modeling the different stages of binaural processing and their influence on perception. Also with respect to the envelope, the question arises to what extent bandpass filtering of the periphery, filtering of the EPSP/IPSP complex (synaptic filter) and possibly downstream weighting by training influence the ITD_{ENV} sensitivity. For this, an experiment similar to the one in Chapter 5 (Klug and Dietz, 2022), with fine frequency spacing, would be important to determine the exact decline across modulation frequency and to deduce the influence of the different stages.

In addition, the experiment in Chapter 5 revealed a steep, cliff-like decline in ITD_{TFS} discrimination between 1300 and 1500 Hz. This finding is in line with previous discrimination experiments (Klumpp and Eady, 1956; Brughera *et al.*, 2013) but contradicting the experiment from Yost (1981) which reports that subjects experienced the full intracranial range of lateralization percepts for IPD-stimuli at frequencies as high as 1.5 kHz. However, this experiment used different paradigms, e.g. longer ramp times and a presentation of a spatial reference between presentations of the target stimuli, than the previously mentioned ones. The influence of these factors needs to be clarified and it needs to be examined whether discrimination and lateralization data can be reconciled. Therefore, an experiment with the same subjects and paradigms is needed to determine discrimination and lateralization.

The steep decline poses a challenge to the current models of binaural processing. The challenge will be to reconcile the various possible sources of decline with behavioural data. Furthermore it would be helpful to measure rate-IPD functions of MSO cells with CFs in the area of the decline. This could provide an important insight into the processing stage at which the steep decline is present. Another focus should be on investigating neural responses at the level of the cortex. It has already been shown (e.g., Ross *et al.*, 2007) that the magnitude of cortical responses measured with Electroencephalography (EEG) in response to IPD changes in on-going sinusoidal stimuli diminishes as the frequency of the tone approaches the upper limit of IPD sensitivity. It might be useful to examine the slope of the decline in EEG-response magnitudes across frequency and possibly measure it with a fine frequency resolution.

7. General conclusion

Our tests have shown a strong effect of hearing loss and age on the decline of IPD_{TFS} (Heinermann *et al.*, 2019). Such results and their modeling could play an important role in individual diagnosis of hearing deficits. An important question is to what extent synaptopathy degrades binaural hearing abilities. The question is whether a reduced number of AN fibers or a change in the properties of the binaural neurons themselves is a better explanation, e.g. for a flattened sensitivity decline.

The models developed in this thesis, with their strong support in physiology and psychoacoustics, are a crucial step towards answering this question.

7.3. References

Bernstein, L. R., and Trahiotis, C. (**2008**). "Discrimination of interaural temporal dispari-
ties conveyed by high-frequency sinusoidally amplitude-modulated tones and high-frequency
transposed tones: Effects of spectrally flanking noises," The Journal of the Acoustical Society
of America **124**(5), 3088–3094, doi: `10.1121/1.2980523`.

Brughera, A., Dunai, L., and Hartmann, W. M. (**2013**). "Human interaural time difference
thresholds for sine tones: The high-frequency limit," J. Acoust. Soc. Am. **133**(5), 2839–2855,
doi: `10.1121/1.4795778`.

Dietz, M., Bernstein, L. R., Trahiotis, C., Ewert, S. D., and Hohmann, V. (**2013**). "The effect
of overall level on sensitivity to interaural differences of time and level at high frequencies,"
The Journal of the Acoustical Society of America **134**(1), 494–502, doi: `10.1121/1.4807827`.

Dreyer, A. A., and Oxenham, A. J. (**2008**). "Effects of level and background noise on interaural
time difference discrimination for transposed stimuli," The Journal of the Acoustical Society
of America **123**(1), EL1–EL7, doi: `10.1121/1.2820442`.

Fredelake, S., and Hohmann, V. (**2012**). "Factors affecting predicted speech intelligibility with
cochlear implants in an auditory model for electrical stimulation," Hearing Research **287**(1),
76–90, doi: `10.1016/j.heares.2012.03.005`.

Furukawa, S. (**2008**). "Detection of combined changes in interaural time and intensity differ-
ences: Segregated mechanisms in cue type and in operating frequency range?," The Journal
of the Acoustical Society of America **123**(3), 1602–1617, doi: `10.1121/1.2835226`.

Hamacher, V. (**2004**). "Signalverarbeitungsmodelle des elektrisch stimulierten Gehörs," Phd
thesis, RWTH Aachen.

Heinermann, H. T., Klug, J., Herrmann, S., Ashida, G., Encke, J., and Dietz, M. (**2019**). "On
the frequency limit of interaural time difference sensitivity for pure tones," in *23. International
Congress on Acoustics*, Aachen, Germany, doi: `10.18154/RWTH-CONV-239106`.

Hu, H., Ausili, S. A., Williges, B., Klug, J., Felsheim, R. C., Vickers, D., and Dietz, M. (**2023**).
"A model framework for simulating spatial hearing of bilateral cochlear implant users," Acta
Acust. **7**, 42, doi: `10.1051/aacus/2023036`.

Hu, H., Klug, J., and Dietz, M. (**2022**). "Simulation of ITD-dependent single-neuron re-
sponses under electrical stimulation and with amplitude-modulated acoustic stimuli," Jour-
nal of the Association for Research in Otolaryngology **23**(4), 535–550, doi: `10.1007/
s10162-021-00823-1`.

Johannesen, P. T., Leclère, T., Wijetillake, A., Segovia-Martínez, M., and Lopez-Poveda, E. A.
(**2022**). "Modeling temporal information encoding by the population of fibers in the healthy
and synaptopathic auditory nerve," Hearing Research **426**, 108621, doi: `10.1016/j.heares.
2022.108621`.

Joris, P. X., and Verschooten, E. (**2013**). "On the limit of neural phase locking to fine structure
in humans," in *Basic Aspects of Hearing*, edited by B. C. J. Moore, R. D. Patterson, I. M.

7. General conclusion

Winter, R. P. Carlyon, and H. E. Gockel, Springer New York, New York, NY, pp. 101–108, doi: `10.1007/978-1-4614-1590-9_12`.

Joris, P. X., and Yin, T. C. (**1992**). "Responses to amplitude-modulated tones in the auditory nerve of the cat," The Journal of the Acoustical Society of America **91**(1), 215–232, doi: `10.1121/1.402757`.

Klug, J., and Dietz, M. (**2022**). "Frequency dependence of sensitivity to interaural phase differences in pure tones," The Journal of the Acoustical Society of America **152**(6), 3130–3141, doi: `10.1121/10.0015246`.

Klug, J., Encke, J., and Dietz, M. (**2023**). "Characterization of the decline in the auditory nerve phase locking at high frequencies," JASA Express Letters **3**(7), 074403, doi: `10.1121/10.0020267`.

Klug, J., Heeringa, A., Köppl, C., and Dietz, M. (**2024**). "Notched noise improves on-frequency ITD$_{\text{ENV}}$ encoding and AN phase locking," in preperation .

Klug, J., Schmors, L., Ashida, G., and Dietz, M. (**2020**). "Neural rate difference model can account for lateralization of high-frequency stimuli," The Journal of the Acoustical Society of America **148**(2), 678–691, doi: `10.1121/10.0001602`.

Klumpp, R. G., and Eady, H. R. (**1956**). "Some measurements of interaural time difference thresholds," J. Acoust. Soc. Am. **28**(5), 859–860, doi: `10.1121/1.1908493`.

Ross, B., Tremblay, K. L., and Picton, T. W. (**2007**). "Physiological detection of interaural phase differences," The Journal of the Acoustical Society of America **121**(2), 1017–1027, doi: `10.1121/1.2404915`.

Verschooten, E., Robles, L., and Joris, P. X. (**2015**). "Assessment of the limits of neural phase-locking using mass potentials," Journal of Neuroscience **35**(5), 2255–2268, doi: `10.1523/JNEUROSCI.2979-14.2015`.

Verschooten, E., Shamma, S., Oxenham, A., Moore, B., Joris, P., Heinz, M., and Plack, C. (**2019**). "The upper frequency limit for the use of phase locking to code temporal fine structure in humans: A compilation of viewpoints," Hearing Research **377**, doi: `10.1016/j.heares.2019.03.011`.

Weiss, T., and Rose, C. (**1988**). "A comparison of synchronization filters in different auditory receptor organs," Hearing Research **33**(2), 175–179, doi: `10.1016/0378-5955(88)90030-5`.

Williges, B., Jürgens, T., Hu, H., and Dietz, M. (**2018**). "Coherent coding of enhanced interaural cues improves sound localization in noise with bilateral cochlear implants," Trends in Hearing **22**, 2331216518781746, doi: `10.1177/2331216518781746` pMID: 29956589.

Yost, W. A. (**1981**). "Lateral position of sinusoids presented with interaural intensive and temporal differences," The Journal of the Acoustical Society of America **70**(2), 397–409, doi: `10.1121/1.386775`.

A. Appendix

This appendix provides additional figures for the various chapters. Most of the figures have been used for illustrative purposes in talks and poster sessions at conferences and are therefore familiar to many researchers.

A.1. Appendix to Chapter 3
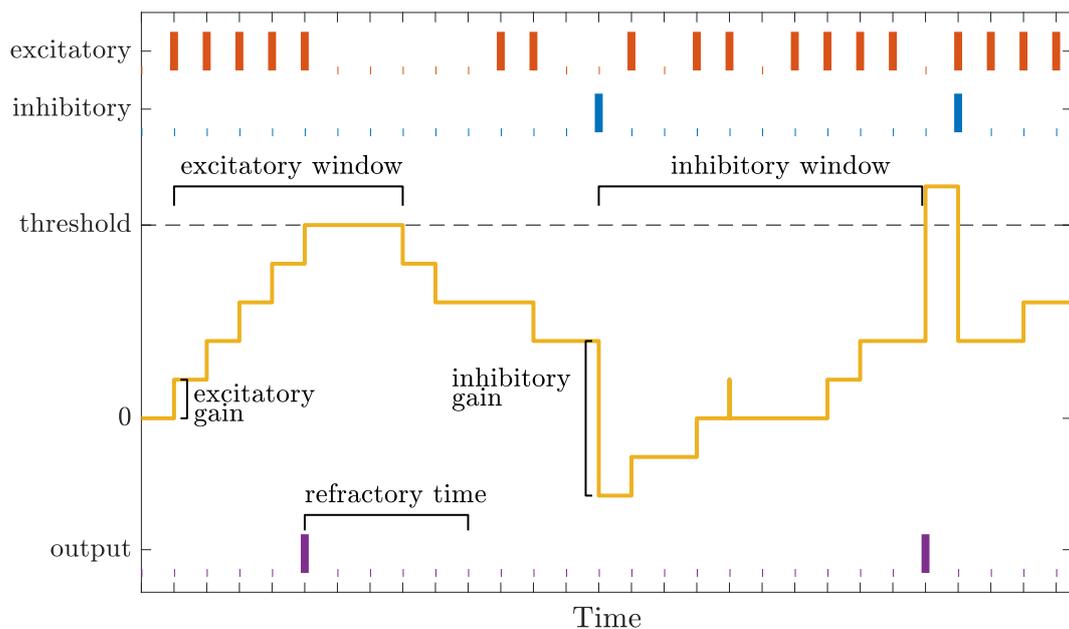
Figure A.1.: Conceptual representation of the EI-model neuron developed by Ashida *et al.* (2016). Red bars in the first row mark the time when an excitatory spike is at the input to the model neuron. Blue bars in the second row mark the time when an inhibitory spike is at the input. The yellow line describes the course of the internal variable (can be understood as volts) across time. Each excitatory input spike increases the internal variable by the *excitatory gain* for the duration of the *excitatory window*. If the internal variable reaches the *threshold* an output spike is generated. The purple bars in the last row mark the time when an output spike is generated. For the duration of the *refractory time*, no further spikes could be generated. Each inhibitory input spike decreases the internal variable by the *inhibitory gain* for the duration of the *inhibitory window*.

A.2. Appendix to Chapter 5

In the example data in Figure A.2 A the 79 %-correct threshold for a 1400-Hz pure-tone is at $\Delta\text{ITD} = 133\,\mu\text{s}$ ($\Delta\text{IPD} = 0.37\pi$). If the sensitivity decline would mirror the decline of AN phase-locking, i.e. decline similarly to a 5th-order low-pass filter (Figure A.2 D), this would result in a correct rate of 72 % at 1500 Hz and 67 % at 1600 Hz for the same ΔIPD. Figure A.2 B shows the psychometric functions determined by the filter order in panel D. A different set of psychometric functions (cf. Figure A.2 C+E) still accounts for the 79 %-correct thresholds up to 1400 Hz but results in correct rate of only 58 % and 52 % for 1500 and 1600 Hz, respectively. This illustrates the critical role of the ITD_{TFS} sensitivity above 1400 Hz for determining the steepness of the sensitivity function.
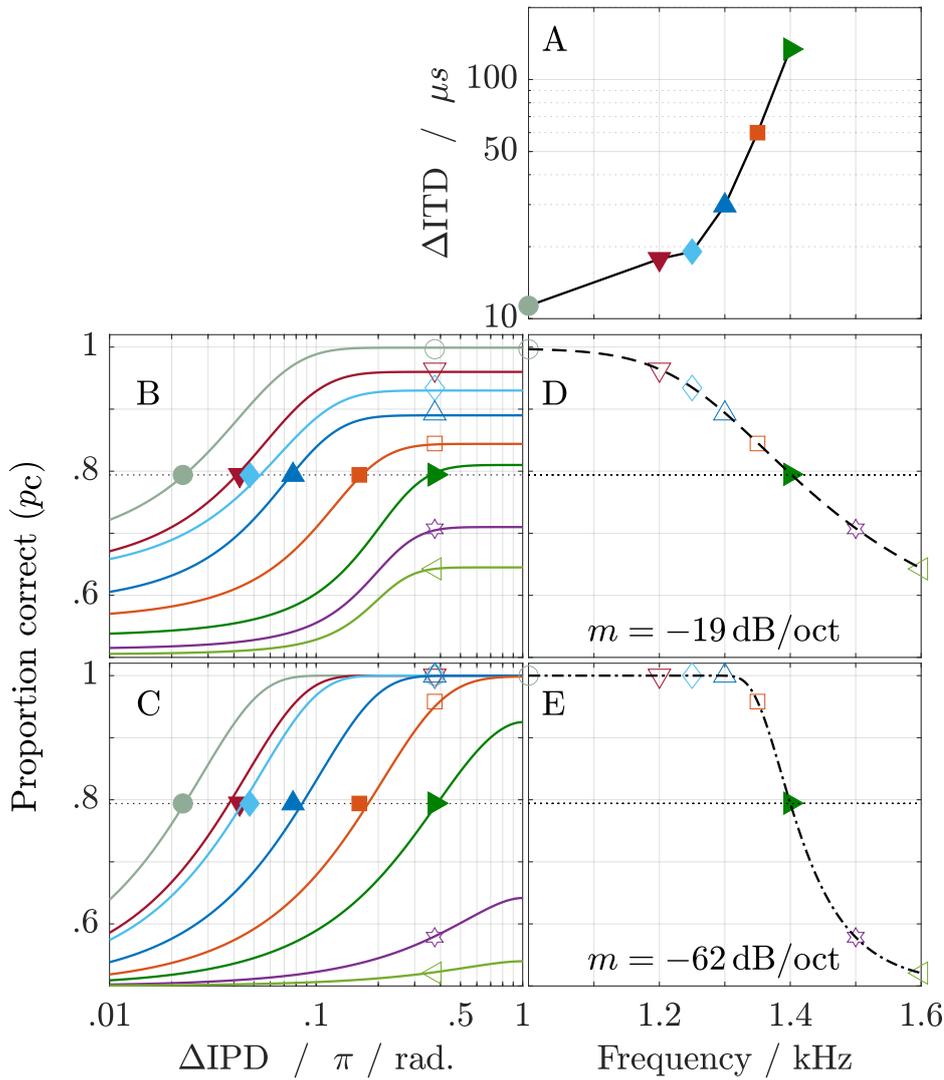
Figure A.2.: **A**: Threshold ΔITDs from (Brughera *et al.*, 2013, Listener 1). **B+C**: Filled symbols indicate the same 79 %-correct threshold as in A together with two different sets of hypothetical psychometric functions. **D+E**: Resulting correct rate as a function of frequency for the two different sets of hypothetical functions at a fixed ΔIPD that corresponds to the experimental threshold $\Delta\text{IPD} = 0.37\pi$ at 1400 Hz.

Bibliography

Altoè, A., Pulkki, V., and Verhulst, S. (**2018**). "The effects of the activation of the inner-hair-cell basolateral K+ channels on auditory nerve responses," Hearing Research **364**, 68–80, doi: `10.1016/j.heares.2018.03.029`.

Ashida, G., Funabiki, K., and Carr, C. (**2013**). "Biophysical basis of the sound analog membrane potential that underlies coincidence detection in the barn owl," Frontiers in Computational Neuroscience **7**, doi: `10.3389/fncom.2013.00102`.

Ashida, G., Kretzberg, J., and Tollin, D. J. (**2016**). "Roles for coincidence detection in coding amplitude-modulated sounds," PLOS Computational Biology **12**(6), 1–27, doi: `10.1371/journal.pcbi.1004997`.

Ashida, G., Tollin, D. J., and Kretzberg, J. (**2017**). "Physiological models of the lateral superior olive," PLOS Computational Biology **13**(12), 1–50, doi: `10.1371/journal.pcbi.1005903`.

Ashida, G., Wagner, H., and Carr, C. E. (**2010**). "Processing of phase-locked spikes and periodic signals," in *Analysis of Parallel Spike Trains*, edited by S. Grün and S. Rotter (Springer US, Boston, MA), pp. 59–74, doi: `10.1007/978-1-4419-5675-0_4`.

Ashmore, J. F. (**1987**). "A fast motile response in guinea-pig outer hair cells: the cellular basis of the cochlear amplifier.," The Journal of Physiology **388**(1), 323–347, doi: `10.1113/jphysiol.1987.sp016617`.

Bernstein, L. R., and Trahiotis, C. (**1996**a). "The normalized correlation: Accounting for binaural detection across center frequency," The Journal of the Acoustical Society of America **100**(6), 3774–3784, doi: `10.1121/1.417237`.

Bernstein, L. R., and Trahiotis, C. (**1996**b). "The normalized correlation: Accounting for binaural detection across center frequency," J. Acoust. Soc. Am. **100**(6), 3774–3784, doi: `10.1121/1.417237`.

Bernstein, L. R., and Trahiotis, C. (**2002**). "Enhancing sensitivity to interaural delays at high frequencies by using "transposed stimuli"," J. Acoust. Soc. Am. **112**(3), 1026–1036, doi: `10.1121/1.1497620`.

Bernstein, L. R., and Trahiotis, C. (**2003**). "Enhancing interaural-delay-based extents of laterality at high frequencies by using 'transposed stimuli'," The Journal of the Acoustical Society of America **113**(6), 3335–3347, doi: `10.1121/1.1570431`.

Bernstein, L. R., and Trahiotis, C. (**2008**). "Discrimination of interaural temporal disparities conveyed by high-frequency sinusoidally amplitude-modulated tones and high-frequency transposed tones: Effects of spectrally flanking noises," The Journal of the Acoustical Society of America **124**(5), 3088–3094, doi: `10.1121/1.2980523`.

Bibliography

Bernstein, L. R., and Trahiotis, C. (**2009**). "How sensitivity to ongoing interaural temporal disparities is affected by manipulations of temporal features of the envelopes of high-frequency stimuli," The Journal of the Acoustical Society of America **125**(5), 3234–3242, doi: `10.1121/1.3101454`.

Bernstein, L. R., and Trahiotis, C. (**2012**). "Lateralization produced by interaural temporal and intensitive disparities of high-frequency, raised-sine stimuli: Data and modeling," The Journal of the Acoustical Society of America **131**(1), 409–415, doi: `10.1121/1.3662056`.

Bernstein, L. R., and Trahiotis, C. (**2016**). "Behavioral manifestations of audiometrically-defined "slight" or "hidden" hearing loss revealed by measures of binaural detection," J. Acoust. Soc. Am. **140**(5), 3540–3548, doi: `10.1121/1.4966113`.

Bernstein, L. R., and Trahiotis, C. (**2017**). "An interaural-correlation-based approach that accounts for a wide variety of binaural detection data," The Journal of the Acoustical Society of America **141**(2), 1150–1160, doi: `10.1121/1.4976098`.

Blauert, J. (**1996**). *Spatial Hearing: The Psychophysics of Human Sound Localization* (The MIT Press).

Boudreau, J. C., and Tsuchitani, C. (**1968**). "Binaural interaction in the cat superior olive s segment.," Journal of Neurophysiology **31**(3), 442–454, doi: `10.1152/jn.1968.31.3.442` pMID: 5687764.

Bouse, J., Vencovský, V., Rund, F., and Marsalek, P. (**2019**). "Functional rate-code models of the auditory brainstem for predicting lateralization and discrimination data of human binaural perception," J. Acoust. Soc. Am. **145**(1), 1–15, doi: `10.1121/1.5084264`.

Brand, A., Behrend, O., Marquardt, T., McAlpine, D., and Grothe, B. (**2002**). "Precise inhibition is essential for microsecond interaural time difference coding," Nature **417**(6888), 543–547, doi: `10.1038/417543a`.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001a**). "Binaural processing model based on contralateral inhibition. i. model structure," J. Acoust. Soc. Am. **110**(2), 1074–1088, doi: `10.1121/1.1383297`.

Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001b**). "Binaural processing model based on contralateral inhibition. ii. dependence on spectral parameters," J. Acoust. Soc. Am. **110**(2), 1089–1104, doi: `10.1121/1.1383298`.

Bruce, I. C., Erfani, Y., and Zilany, M. S. (**2018**). "A phenomenological model of the synapse between the inner hair cell and auditory nerve: Implications of limited neurotransmitter release sites," Hearing Research **360**, 40–54, doi: `10.1016/j.heares.2017.12.016` computational models of the auditory system.

Brughera, A., Dunai, L., and Hartmann, W. M. (**2013**). "Human interaural time difference thresholds for sine tones: The high-frequency limit," J. Acoust. Soc. Am. **133**(5), 2839–2855, doi: `10.1121/1.4795778`.

Cai, H., Carney, L. H., and Colburn, H. S. (**1998**). "A model for binaural response properties of inferior colliculus neurons. ii. a model with interaural time difference-sensitive excitatory and inhibitory inputs and an adaptation mechanism," The Journal of the Acoustical Society of America **103**(1), 494–506, doi: `10.1121/1.421130`.

Cant, N. B., and Casseday, J. H. (**1986**). "Projections from the anteroventral cochlear nucleus to the lateral and medial superior olivary nuclei," Journal of Comparative Neurology **247**(4), 457–476, doi: `10.1002/cne.902470406`.

Cant, N. B., and Hyson, R. L. (**1992**). "Projections from the lateral nucleus of the trapezoid body to the medial superior olivary nucleus in the gerbil," Hearing research **58**(1), 26–34, doi: `10.1016/0378-5955(92)90005-8`.

Carr, C., and Konishi, M. (**1990**). "A circuit for detection of interaural time differences in the brain stem of the barn owl," Journal of Neuroscience **10**(10), 3227–3246, doi: `10.1523/JNEUROSCI.10-10-03227.1990`.

Carr, C. E., and Konishi, M. (**1988**). "Axonal delay lines for time measurement in the owl's brainstem," Proceedings of the National Academy of Sciences **85**(21), 8311–8315, doi: `10.1073/pnas.85.21.8311`.

Colburn, H. S. (**1973**). "Theory of binaural interaction based on auditory-nerve data. i. general strategy and preliminary results on interaural discrimination," The Journal of the Acoustical Society of America **54**(6), 1458–1470, doi: `10.1121/1.1914445`.

Colburn, H. S. (**1977**). "Theory of binaural interaction based on auditory-nerve data. ii. detection of tones in noise," The Journal of the Acoustical Society of America **61**(2), 525–533, doi: `10.1121/1.381294`.

Colburn, H. S., and Durlach, N. I. (**1978**). "Models of binaural interaction," in *Handbook of perception: Hearing*, edited by E. C. Carterette and M. P. Friedman, **4** (Academic Press, New York), pp. 467–518.

Colburn, H. S., and Esquissaud, P. (**1976**). "An auditroy-nerve model for interaural time discrimination of high-freqeuncy complex stimuli," The Journal of the Acoustical Society of America **59**(S1), S23–S23, doi: `10.1121/1.2002503`.

Colburn, H. S., Yan-an, H., and Culotta, C. P. (**1990**). "Coincidence model of mso responses," Hearing research **49**(1), 335–346, doi: `10.1016/0378-5955(90)90112-3`.

Culling, J. F. (**2007**). "Evidence specifically favoring the equalization-cancellation theory of binaural unmasking," The Journal of the Acoustical Society of America **122**(5), 2803–2813, doi: `10.1121/1.2785035`.

Dau, T., Püschel, D., and Kohlrausch, A. (**1996**). "A quantitative model of the "effective" signal processing in the auditory system. i. model structure," The Journal of the Acoustical Society of America **99**(6), 3615–3622, doi: `10.1121/1.414959`.

Dietz, M., Bernstein, L. R., Trahiotis, C., Ewert, S. D., and Hohmann, V. (**2013**). "The effect of overall level on sensitivity to interaural differences of time and level at high frequencies," The Journal of the Acoustical Society of America **134**(1), 494–502, doi: `10.1121/1.4807827`.

Bibliography

Dietz, M., Ewert, S. D., and Hohmann, V. (**2009**). "Lateralization of stimuli with independent fine-structure and envelope-based temporal disparities," J. Acoust. Soc. Am. **125**(3), 1622–1635, doi: `10.1121/1.3076045`.

Dietz, M., Ewert, S. D., and Hohmann, V. (**2011**). "Auditory model based direction estimation of concurrent speakers from binaural signals," Speech Communication **53**(5), 592–605, doi: `10.1016/j.specom.2010.05.006` perceptual and Statistical Audition.

Dietz, M., Klein-Hennig, M., and Hohmann, V. (**2015**). "The influence of pause, attack, and decay duration of the ongoing envelope on sound lateralization," The Journal of the Acoustical Society of America **137**(2), EL137–EL143, doi: `10.1121/1.4905891`.

Dietz, M., Lestang, J.-H., Majdak, P., Stern, R. M., Marquardt, T., Ewert, S. D., Hartmann, W. M., and Goodman, D. F. (**2018**). "A framework for testing and comparing binaural models," Hearing Research **360**, 92–106, doi: `10.1016/j.heares.2017.11.010` computational models of the auditory system.

Dietz, M., Wang, L., Greenberg, D., and McAlpine, D. (**2016**). "Sensitivity to interaural time differences conveyed in the stimulus envelope: Estimating inputs of binaural neurons through the temporal analysis of spike trains," The Journal of the Association for Research in Otolaryngology **17**, 313–330, doi: `10.1007/s10162-016-0573-9`.

Dreyer, A., and Delgutte, B. (**2006**a). "Phase locking of auditory-nerve fibers to the envelopes of high-freqeuncy sounds: Implications for sound localization," Journal of Neurophysiology **96**(5), 2327–2341, doi: `10.1152/jn.00326.2006`.

Dreyer, A., and Delgutte, B. (**2006**b). "Phase locking of auditory-nerve fibers to the envelopes of high-frequency sounds: Implications for sound localization," Journal of Neurophysiology **96**(5), 2327–2341, doi: `10.1152/jn.00326.2006` pMID: 16807349.

Dreyer, A. A., and Oxenham, A. J. (**2008**). "Effects of level and background noise on interaural time difference discrimination for transposed stimuli," The Journal of the Acoustical Society of America **123**(1), EL1–EL7, doi: `10.1121/1.2820442`.

Dye, R. H. (**1990**). "The combination of interaural information across frequencies: Lateralization on the basis of interaural delay," J. Acoust. Soc. Am. **88**(5), 2159–2170, doi: `10.1121/1.400113`.

Encke, J., and Dietz, M. (**2022**). "A hemispheric two-channel code accounts for binaural unmasking in humans," Communications Biology **5**(1), 1122, doi: `10.1038/s42003-022-04098-x`.

Encke, J., and Hemmert, W. (**2018**). "Extraction of inter-aural time differences using a spiking neuron network model of the medial superior olive," Frontiers in Neuroscience **12**, 140, doi: `10.3389/fnins.2018.00140`.

Eurich, B., Encke, J., Ewert, S. D., and Dietz, M. (**2022**). "Lower interaural coherence in off-signal bands impairs binaural detection," The Journal of the Acoustical Society of America **151**(6), 3927–3936, doi: `10.1121/10.0011673`.

Ewert, S. D. (**2013**). "AFC—A modular framework for running psychoacoustic experiments and computational perception models," in *Proceedings of the international conference on acoustics AIA-DAGA*, pp. 1326–1329.

Fisher, N. I. (**1993**). *Statistical Analysis of Circular Data* (Cambridge University Press).

Folkerts, M. L., and Stecker, G. C. (**2022**). "Spectral weighting functions for lateralization and localization of complex sounda)," The Journal of the Acoustical Society of America **151**(5), 3409–3425, doi: `10.1121/10.0011469`.

Fredelake, S., and Hohmann, V. (**2012**). "Factors affecting predicted speech intelligibility with cochlear implants in an auditory model for electrical stimulation," Hearing Research **287**(1), 76–90, doi: `10.1016/j.heares.2012.03.005`.

Furukawa, S. (**2008**). "Detection of combined changes in interaural time and intensity differences: Segregated mechanisms in cue type and in operating frequency range?," The Journal of the Acoustical Society of America **123**(3), 1602–1617, doi: `10.1121/1.2835226`.

Glasberg, B. R., and Moore, B. C. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hearing Research **47**(1), 103–138, doi: `10.1016/0378-5955(90)90170-T`.

Goldberg, J. M., and Brown, P. B. (**1969**). "Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization.," Journal of Neurophysiology **32**(4), 613–636, doi: `10.1152/jn.1969.32.4.613`.

Goodman, D. F., Benichoux, V., and Brette, R. (**2013**). "Decoding neural responses to temporal cues for sound localization," eLife **2**, doi: `10.7554/eLife.01312.001`.

Goupell, M. J., Stecker, G. C., and Tollin, D. J. (**2023**). "Is the rapid decline in interaural time difference sensitivity above 700 Hz explained by downward spread of excitation into the frequency dominant region?," The Journal of the Acoustical Society of America **153**(3_supplement), A335–A335, `https://www.researchgate.net/publication/370307645_Is_the_rapid_decline_in_interaural_time_difference_sensitivity_above_700_Hz_explained_by_downward_spread_of_excitation_into_the_frequency_dominant_region`, doi: `10.1121/10.0019056`.

Green, D. M., and Swets, J. A. (**1966**). *Signal Detection Theory and Psychophysics* (Wiley, New York).

Greenberg, D., Monaghan, J. J. M., Dietz, M., Marquardt, T., and McAlpine, D. (**2017**). "Influence of envelope waveform on itd sensitivity of neurons in the auditory midbrain," Journal of Neurophysiology **118**(4), 2358–2370, doi: `10.1152/jn.01048.2015`.

Greenwood, D. D. (**1961**). "Critical bandwidth and the frequency coordinates of the basilar membrane," The Journal of the Acoustical Society of America **33**(10), 1344–1356, doi: `10.1121/1.1908437`.

Grothe, B., Pecka, M., and McAlpine, D. (**2010**). "Mechanisms of sound localization in mammals," Physiological Reviews **90**(3), 983–1012, doi: `10.1152/physrev.00026.2009`.

Bibliography

Grothe, B., and Sanes, D. H. (**1993**). "Bilateral inhibition by glycinergic afferents in the medial superior olive," Journal of Neurophysiology **69**(4), 1192–1196, doi: `10.1152/jn.1993.69.4.1192` pMID: 8492158.

Hafter, E. R., Dye, R. H., and Gilkey, R. H. (**1979**). "Lateralization of tonal signals which have neither onsets nor offsets," J. Acoust. Soc. Am. **65**(2), 471–477, doi: `10.1121/1.382346`.

Hamacher, V. (**2004**). "Signalverarbeitungsmodelle des elektrisch stimulierten Gehörs," Phd thesis, RWTH Aachen.

Heeringa, A. N., Zhang, L., Ashida, G., Beutelmann, R., Steenken, F., and Köppl, C. (**2020**). "Temporal coding of single auditory nerve fibers is not degraded in aging gerbils," Journal of Neuroscience **40**(2), 343–354, doi: `10.1523/JNEUROSCI.2784-18.2019`.

Heil, P., Neubauer, H., and Irvine, D. R. F. (**2011**). "An improved model for the rate–level functions of auditory-nerve fibers," Journal of Neuroscience **31**(43), 15424–15437, doi: `10.1523/JNEUROSCI.1638-11.2011`.

Heinermann, H. T., Klug, J., Herrmann, S., Ashida, G., Encke, J., and Dietz, M. (**2019**). "On the frequency limit of interaural time difference sensitivity for pure tones," in *23. International Congress on Acoustics*, Aachen, Germany, doi: `10.18154/RWTH-CONV-239106`.

Heinz, M. G., Colburn, H. S., and Carney, L. H. (**2001**). "Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the Auditory Nerve," Neural Computation **13**(10), 2273–2316, doi: `10.1162/089976601750541804`.

Henning, G. (**1983**). "Lateralization of low-frequency transients," Hearing Research **9**(2), 153–172, doi: `10.1016/0378-5955(83)90025-4`.

Hohmann, V. (**2002**). "Frequency analysis and synthesis using a gammatone filterbank," Acta Acustica united with Acustica **88**(3), 433–442.

Hu, H., Ausili, S. A., Williges, B., Klug, J., Felsheim, R. C., Vickers, D., and Dietz, M. (**2023**). "A model framework for simulating spatial hearing of bilateral cochlear implant users," Acta Acust. **7**, 42, doi: `10.1051/aacus/2023036`.

Hu, H., Klug, J., and Dietz, M. (**2022**). "Simulation of ITD-dependent single-neuron responses under electrical stimulation and with amplitude-modulated acoustic stimuli," Journal of the Association for Research in Otolaryngology **23**(4), 535–550, doi: `10.1007/s10162-021-00823-1`.

Hudspeth, A. J. (**1985**). "The cellular basis of hearing: The biophysics of hair cells," Science **230**(4727), 745–752, doi: `10.1126/science.2414845`.

Jeffress, L. (**1948**). "A place theory of sound localization," Journal of comparative and physiological psychology **41**(1), 35—39, doi: `10.1037/h0061495`.

Jercog, P. E., Svirskis, G., Kotak, V. C., Sanes, D. H., and Rinzel, J. (**2010**). "Asymmetric excitatory synaptic dynamics underlie interaural time difference processing in the auditory system," PLOS Biology **8**(6), 1–9, doi: `10.1371/journal.pbio.1000406`.

Johannesen, P. T., Leclère, T., Wijetillake, A., Segovia-Martínez, M., and Lopez-Poveda, E. A. (**2022**). "Modeling temporal information encoding by the population of fibers in the healthy and synaptopathic auditory nerve," Hearing Research **426**, 108621, doi: `10.1016/j.heares.2022.108621`.

John, M. S., Dimitrijevic, A., and Picton, T. (**2002**). "Auditory steady-state responses to exponential modulation envelopes," Ear and Hearing **23**, 106–117, doi: `10.1097/00003446-200204000-00004`.

Johnson, D. H. (**1980**). "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," The Journal of the Acoustical Society of America **68**(4), 1115–1122, doi: `10.1121/1.384982`.

Johnson, N. L., Kotz, S., and Kemp, A. W. (**1993**). *Univariate Discrete Distributions* (Wiley-Interscience, Hoboken, NJ).

Joris, P. X. (**1996**). "Envelope coding in the lateral superior olive. ii. characteristic delays and comparison with responses in the medial superior olive," Journal of Neurophysiology **76**(4), 2137–2156, doi: `10.1152/jn.1996.76.4.2137`.

Joris, P. X., Carney, L. H., Smith, P. H., and Yin, T. C. (**1994**). "Enhancement of neural synchronization in the anteroventral cochlear nucleus. i. responses to tones at the characteristic frequency," Journal of Neurophysiology **71**(3), 1022–1036, doi: `10.1152/jn.1994.71.3.1022`.

Joris, P. X., and van der Heijden, M. (**2019**). "Early binaural hearing: The comparison of temporal differences at the two ears," Annual Review of Neuroscience **42**(1), 433–457, doi: `10.1146/annurev-neuro-080317-061925`.

Joris, P. X., and Verschooten, E. (**2013**). "On the limit of neural phase locking to fine structure in humans," in *Basic Aspects of Hearing*, edited by B. C. J. Moore, R. D. Patterson, I. M. Winter, R. P. Carlyon, and H. E. Gockel, Springer New York, New York, NY, pp. 101–108, doi: `10.1007/978-1-4614-1590-9_12`.

Joris, P. X., and Yin, T. C. (**1992**). "Responses to amplitude-modulated tones in the auditory nerve of the cat," The Journal of the Acoustical Society of America **91**(1), 215–232, doi: `10.1121/1.402757`.

Joris, P. X., and Yin, T. C. (**1995**). "Envelope coding in the lateral superior olive. I. Sensitivity to interaural time differences," Journal of Neurophysiology **73**(3), 1043–1062, doi: `10.1152/jn.1995.73.3.1043`.

Joris, P. X., and Yin, T. C. T. (**1998**). "Envelope coding in the lateral superior olive. iii. comparison with afferent pathways," Journal of Neurophysiology **79**(1), 253–269, doi: `10.1152/jn.1998.79.1.253`.

Jürgens, T., and Brand, T. (**2009**). "Microscopic prediction of speech recognition for listeners with normal hearing in noise using an auditory model," The Journal of the Acoustical Society of America **126**(5), 2635–2648, doi: `10.1121/1.3224721`.

Bibliography

Kayser, H., Ewert, S., Anemüller, J., Rohdenburg, T., Hohmann, V., and Kollmeier, B. (**2009**). "Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses," Eurasip Journal on Advances in Signal Processing **2009**, 10, doi: `10.1155/2009/298605`.

Kelvasa, D., and Dietz, M. (**2015**). "Auditory model-based sound direction estimation with bilateral cochlear implants," Trends in Hearing **19**, 2331216515616378, doi: `10.1177/2331216515616378`.

Kiang, N. Y., Watanabe, T., Thomas, C., and Clark, L. (**1965**). *Discharge Patterns of single fibers in the cat's auditory nerve* (M.I.T. Press. Research Monograph 35, Cambridge, MA).

Kil, J., Hkageyama, G., Semple, M. N., and Kitzes, L. M. (**1995**). "Development of ventral cochlear nucleus projections to the superior olivary complex in gerbil," Journal of Comparative Neurology **353**(3), 317–340, doi: `10.1002/cne.903530302`.

Klein-Hennig, M., Dietz, M., Hohmann, V., and Ewert, S. D. (**2011**). "The influence of different segments of the ongoing envelope on sensitivity to interaural time delays," The Journal of the Acoustical Society of America **129**(6), 3856–3872, doi: `10.1121/1.3585847`.

Klug, J., and Dietz, M. (**2022**). "Frequency dependence of sensitivity to interaural phase differences in pure tones," The Journal of the Acoustical Society of America **152**(6), 3130–3141, doi: `10.1121/10.0015246`.

Klug, J., Encke, J., and Dietz, M. (**2023**). "Characterization of the decline in the auditory nerve phase locking at high frequencies," JASA Express Letters **3**(7), 074403, doi: `10.1121/10.0020267`.

Klug, J., Heeringa, A., Köppl, C., and Dietz, M. (**2024**). "Notched noise improves on-frequency ITD_{ENV} encoding and AN phase locking," in preperation .

Klug, J., Schmors, L., Ashida, G., and Dietz, M. (**2020**). "Neural rate difference model can account for lateralization of high-frequency stimuli," The Journal of the Acoustical Society of America **148**(2), 678–691, doi: `10.1121/10.0001602`.

Klumpp, R. G., and Eady, H. R. (**1956**). "Some measurements of interaural time difference thresholds," J. Acoust. Soc. Am. **28**(5), 859–860, doi: `10.1121/1.1908493`.

Kuwabara, N., and Zook, J. M. (**1992**). "Projections to the medial superior olive from the medial and lateral nuclei of the trapezoid body in rodents and bats," Journal of Comparative Neurology **324**(4), 522–538, doi: `10.1002/cne.903240406`.

Köppl, C., and Carr, C. E. (**2008**). "Maps of interaural time difference in the chicken's brainstem nucleus laminaris," Biological Cybernetics **9**, 541–559, doi: `10.1007/s00422-008-0220-6`.

Leibold, C., and Grothe, B. (**2015**). "Sound localization with microsecond precision in mammals: What is it we do not understand?," e-Neuroforum **21**(1), 3–10, doi: `doi:10.1515/s13295-015-0001-3`.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**(2B), 467–477, doi: `10.1121/1.1912375`.

Liberman, M. C. (**1978**). "Auditory-nerve response from cats raised in a low-noise chamber," The Journal of the Acoustical Society of America **63**(2), 442–455, doi: `10.1121/1.381736`.

Lindemann, W. (**1986**). "Extension of a binaural cross-correlation model by contralateral inhibition: I. Simulation of lateralization for stationary signals," The Journal of the Acoustical Society of America **80**(6), 1608–1622, doi: `10.1121/1.394325`.

Liu, L.-F., Palmer, A. R., and Wallace, M. N. (**2006**). "Phase-locked responses to pure tones in the inferior colliculus," Journal of Neurophysiology **95**(3), 1926–1935, doi: `10.1152/jn.00497.2005`.

Lopez-Poveda, E. A., and Eustaquio-Martín, A. (**2006**). "A biophysical model of the inner hair cell: The contribution of potassium currents to peripheral auditory compression," Journal of the Association for Research in Otolaryngology **7**, 218–235, doi: `10.1007/s10162-006-0037-8`.

Lopez-Poveda, E. A., and Meddis, R. (**2001**). "A human nonlinear cochlear filterbank," The Journal of the Acoustical Society of America **110**(6), 3107–3118, doi: `10.1121/1.1416197`.

Macpherson, E. A., and Middlebrooks, J. C. (**2002**). "Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited," The Journal of the Acoustical Society of America **111**(5), 2219–2236, doi: `10.1121/1.1471898`.

Magezi, D. A., and Krumbholz, K. (**2010**). "Evidence for opponent-channel coding of interaural time differences in human auditory cortex," Journal of Neurophysiology **104**(4), 1997–2007, doi: `10.1152/jn.00424.2009` pMID: 20702739.

Marquardt, T., and Mcalpine, D. (**2007**). "A π-limit for coding ITDs: Implications for binaural models," in *Hearing – From Sensory Processing to Perception*, edited by B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, and J. Verhey, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 407–416, doi: `10.1007/978-3-540-73009-5_44`.

Mayo, P. G., and Goupell, M. J. (**2020**). "Acoustic factors affecting interaural level differences for cochlear-implant users," The Journal of the Acoustical Society of America **147**(4), EL357–EL362, doi: `10.1121/10.0001088`.

McAlpine, D., Jiang, D., and Palmer, A. R. (**2001**). "A neural code for low-frequency sound localization in mammals," Nature neuroscience **4**(4), 396–401, doi: `10.1038/86049`.

Meddis, R., Lopez-Poveda, E. A., Fay, R. R., and Popper, A. N. (**2010**). *Computational Models of the Auditory System* (Springer, New York).

Miller, C. A., Abbas, P. J., and Robinson, B. (**2001**). "Response properties of the refractory auditory nerve fiber," The Journal of the Association for Research in Otolaryngology **2**, 216–232, doi: `10.1007/s101620010083`.

Mills, A. W. (**1958**). "On the minimum audible angle," J. Acoust. Soc. Am. **30**(4), 237–246, doi: `10.1121/1.1909553`.

Bibliography

Monaghan, J. J. M., Bleeck, S., and McAlpine, D. (**2015**). "Sensitivity to envelope interaural time differences at high modulation rates," Trends in Hearing **19**, 2331216515619331, doi: `10.1177/2331216515619331`.

Moore, B. C. (**2021**). "Effects of hearing loss and age on the binaural processing of temporal envelope and temporal fine structure information," Hearing Research **402**, 107991, doi: `10.1016/j.heares.2020.107991` special Issue on Presbycusis.

Moore, B. J. C. (**2008**). "The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people," Journal of the Association for Research in Otolaryngology **9**, 399–406, doi: `10.1007/s10162-008-0143-x`.

Myoga, M. H., Lehnert, S., Leibold, C., Felmy, F., and Grothe, B. (**2014**). "Glycinergic inhibition tunes coincidence detection in the auditory brainstem," Nature Communications **5**, doi: `10.1038/ncomms4790`.

Osen, K. K. (**1969**). "The intrinsic organization of the cochlear nuclei in the cat," Acta Oto-Laryngologica **67**(2-6), 352–359, doi: `10.3109/00016486909125462` pMID: 5374653.

Osses Vecchi, A., Varnet, L., Carney, L. H., Dau, T., Bruce, I. C., Verhulst, S., and Majdak, P. (**2022**). "A comparative study of eight human auditory models of monaural processing," Acta Acust. **6**, 17, doi: `10.1051/aacus/2022008`.

Owrutsky, Z. L., Benichoux, V., and Tollin, D. J. (**2021**). "Binaural hearing by the mammalian auditory brainstem: Joint coding of interaural level and time differences by the lateral superior olive," in *Binaural Hearing*, edited by R. Y. Litovsky, M. J. Goupell, R. R. Fay, and A. N. Popper (Springer International Publishing, Cham), pp. 113–144, doi: `10.1007/978-3-030-57100-9_5`.

Palmer, A. R., and Russell, I. J. (**1986**). "Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells," Hearing Research **24**(1), 1–15, doi: `10.1016/0378-5955(86)90002-X`.

Pavão, R., Sussman, E. S., Fischer, B. J., and Peña, J. L. (**2020**). "Natural ITD statistics predict human auditory spatial perception," Elife **9**, doi: `10.7554/eLife.51927`.

Pecka, M., Brand, A., Behrend, O., and Grothe, B. (**2008**). "Interaural time difference processing in the mammalian medial superior olive: The role of glycinergic inhibition," Journal of Neuroscience **28**(27), 6914–6925, doi: `10.1523/JNEUROSCI.1660-08.2008`.

Peterson, A. J., and Heil, P. (**2020**). "Phase locking of auditory nerve fibers: The role of lowpass filtering by hair cells," Journal of Neuroscience **40**(24), 4700–4714, doi: `10.1523/JNEUROSCI.2269-19.2020`.

Pickles, J. (**2015**). "Auditory pathways: Anatomy and physiology," Handbook of clinical neurology **129C**, 3–25, doi: `10.1016/B978-0-444-62630-1.00001-9`.

Plack, C. (**2013**). *The sense of hearing: Second Edition* (Psychology Press, New York).

Remme, M. W. H., Donato, R., Mikiel-Hunter, J., Ballestero, J. A., Foster, S., Rinzel, J., and McAlpine, D. (**2014**). "Subthreshold resonance properties contribute to the efficient coding of auditory spatial cues," Proceedings of the National Academy of Sciences **111**(22), E2339–E2348, doi: `10.1073/pnas.1316216111`.

Roberts, M. T., Seeman, S. C., and Golding, N. L. (**2013**). "A mechanistic understanding of the role of feedforward inhibition in the mammalian sound localization circuitry," Neuron **78**(5), 923–935, doi: `10.1016/j.neuron.2013.04.022`.

Robles, L., and Ruggero, M. A. (**2001**). "Mechanics of the mammalian cochlea," Physiological Reviews **81**(3), 1305–1352, doi: `10.1152/physrev.2001.81.3.1305`.

Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (**1967**). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey.," Journal of Neurophysiology **30**(4), 769–793, doi: `10.1152/jn.1967.30.4.769`.

Ross, B., Tremblay, K. L., and Picton, T. W. (**2007**). "Physiological detection of interaural phase differences," The Journal of the Acoustical Society of America **121**(2), 1017–1027, doi: `10.1121/1.2404915`.

Sanes, D. (**1990**). "An in vitro analysis of sound localization mechanisms in the gerbil lateral superior olive," Journal of Neuroscience **10**(11), 3494–3506, doi: `10.1523/JNEUROSCI.10-11-03494.1990`.

Sayers, B. M. (**1964**). "Acoustic-image lateralization judgments with binaural tones," J. Acoust. Soc. Am. **36**(5), 923–926, doi: `10.1121/1.1919121`.

Schaette, R., and McAlpine, D. (**2011**). "Tinnitus with a normal audiogram: Physiological evidence for hidden hearing loss and computational model," **31**(38), 13452–13457, doi: `10.1523/JNEUROSCI.2156-11.2011`.

Schwartz, I. R. (**1992**). "The superior olivary complex and lateral lemniscal nuclei," in *The Mammalian Auditory Pathway: Neuroanatomy*, edited by D. B. Webster, A. N. Popper, and R. R. Fay (Springer New York, New York, NY), pp. 117–167, doi: `10.1007/978-1-4612-4416-5_4`.

Silberberg, G., Wu, C., and Markram, H. (**2004**). "Synaptic dynamics control the timing of neuronal excitation in the activated neocortical microcircuit," The Journal of Physiology **556**(1), 19–27, doi: `10.1113/jphysiol.2004.060962`.

Smith, P. H., Joris, P. X., and Yin, T. C. T. (**1993**). "Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: Evidence for delay lines to the medial superior olive," Journal of Comparative Neurology **331**(2), 245–260, doi: `10.1002/cne.903310208`.

Søndergaard, P. L., and Majdak, P. (**2013**). "The auditory modeling toolbox," in *The Technology of Binaural Listening*, edited by J. Blauert (Springer Berlin Heidelberg, Berlin, Heidelberg), pp. 33–56, `https://www.amtoolbox.org/`.

Bibliography

Spangler, K. M., Warr, W. B., and Henkel, C. K. (**1985**). "The projections of principal cells of the medial nucleus of the trapezoid body in the cat," Journal of Comparative Neurology **238**(3), 249–262, doi: `h10.1002/cne.902380302`.

Stecker, G. C., Harrington, I. A., and Middlebrooks, J. C. (**2005**). "Location coding by opponent neural populations in the auditory cortex," PLOS Biology **3**(3), doi: `10.1371/journal.pbio.0030078`.

Stern, R. M., and Colburn, H. S. (**1978**). "Theory of binaural interaction based on auditory-nerve data. IV. A model for subjective lateral position," J. Acoust. Soc. Am. **64**(1), 127–140, doi: `10.1121/1.381978`.

Stern, R. M., and Shear, G. D. (**1996**a). "Lateralization and detection of low-freqeuncy binaural stimuli: Effects of distribution of internal delay," The Journal of the Acoustical Society of America **100**(4), 2278–2288, doi: `10.1121/1.417937`.

Stern, R. M., and Shear, G. D. (**1996**b). "Lateralization and detection of low-frequency binaural stimuli: Effects of distribution of internal delay," J. Acoust. Soc. Am. **100**(4), 2278–2288, doi: `10.1121/1.417937`.

Sterratt, D., Graham, B., Gillies, A., and Willshaw, D. (**2011**). *Principles of Computational Modelling in Neuroscience* (Cambridge University Press).

Stotler, W. A. (**1953**). "An experimental study of the cells and connections of the superior olivary complex of the cat," Journal of Comparative Neurology **98**(3), 401–431, doi: `10.1002/cne.900980303`.

Strutt, J. W. (**1907**). "XII. On our perception of sound direction," The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **13**(74), 214–232, doi: `10.1080/14786440709463595`.

Svirskis, G., Kotak, V., Sanes, D. H., and Rinzel, J. (**2004**). "Sodium along with low-threshold potassium currents enhance coincidence detection of subthreshold noisy signals in mso neurons," Journal of Neurophysiology **91**(6), 2465–2473, doi: `10.1152/jn.00717.2003` pMID: 14749317.

Takanen, M., Santala, O., and Pulkki, V. (**2014**). "Visualization of functional count-comparison-based binaural auditory model output," Hearing Research **309**, 147–163, doi: `10.1016/j.heares.2013.10.004`.

Thavam, S., and Dietz, M. (**2019**). "Smallest perceivable interaural time differences," J. Acoust. Soc. Am. **145**(1), 458–468, doi: `10.1121/1.5087566`.

Tollin, D. J. (**2003**). "The lateral superior olive: A functional role in sound source localization," The Neuroscientist **9**(2), 127–143, doi: `10.1177/1073858403252228` pMID: 12708617.

Tollin, D. J., and Yin, T. C. T. (**2002**a). "The coding of spatial location by single units in the lateral superior olive of the cat. I. Spatial receptive fields in azimuth," Journal of Neuroscience **22**(4), 1454–1467, doi: `10.1523/JNEUROSCI.22-04-01454.2002`.

120

Tollin, D. J., and Yin, T. C. T. (**2002**b). "The coding of spatial location by single units in the lateral superior olive of the cat. i. spatial receptive fields in azimuth," **22**(4), 1454–1467, doi: `10.1523/JNEUROSCI.22-04-01454.2002`.

Tsai, J. J., Koka, K., and Tollin, D. J. (**2010**). "Varying overall sound intensity to the two ears impacts interaural level difference discrimination thresholds by single neurons in the lateral superior olive," Journal of Neurophysiology **103**(2), 875–886, doi: `10.1152/jn.00911.2009`.

Tsuchitani, C. (**1988**). "The inhibition of cat lateral superior olive unit excitatory responses to binaural tone bursts. II. The sustained discharges," Journal of Neurophysiology **59**(1), 184–211, doi: `10.1152/jn.1988.59.1.184`.

van de Par, S., and Kohlrausch, A. (**1997**). "A new approach to comparing binaural masking level differences at low and high frequencies," The Journal of the Acoustical Society of America **101**(3), 1671–1680, doi: `10.1121/1.418151`.

Verhulst, S., Bharadwaj, H. M., Mehraei, G., Shera, C. A., and Shinn-Cunningham, B. G. (**2015**). "Functional modeling of the human auditory brainstem response to broadband stimulation," The Journal of the Acoustical Society of America **138**(3), 1637–1659, doi: `10.1121/1.4928305`.

Verhulst, S., Dau, T., and Shera, C. A. (**2012**). "Nonlinear time-domain cochlear model for transient stimulation and human otoacoustic emission," The Journal of the Acoustical Society of America **132**(6), 3842–3848, doi: `10.1121/1.4763989`.

Verschooten, E., Robles, L., and Joris, P. X. (**2015**). "Assessment of the limits of neural phase-locking using mass potentials," Journal of Neuroscience **35**(5), 2255–2268, doi: `10.1523/JNEUROSCI.2979-14.2015`.

Verschooten, E., Shamma, S., Oxenham, A., Moore, B., Joris, P., Heinz, M., and Plack, C. (**2019**). "The upper frequency limit for the use of phase locking to code temporal fine structure in humans: A compilation of viewpoints," Hearing Research **377**, doi: `10.1016/j.heares.2019.03.011`.

von Békésy, G. (**1970**). "Travelling waves as frequency analysers in the cochlea," Nature **225**, 1207–1209, doi: `10.1038/2251207a0`.

Weiss, T., and Rose, C. (**1988**a). "A comparison of synchronization filters in different auditory receptor organs," Hearing Research **33**(2), 175–179, doi: `10.1016/0378-5955(88)90030-5`.

Weiss, T., and Rose, C. (**1988**b). "Stages of degradation of timing information in the cochlea: A comparison of hair-cell and nerve-fiber responses in the alligator lizard," Hearing Research **33**(2), 167–174, doi: `10.1016/0378-5955(88)90029-9`.

Williges, B., Jürgens, T., Hu, H., and Dietz, M. (**2018**). "Coherent coding of enhanced interaural cues improves sound localization in noise with bilateral cochlear implants," Trends in Hearing **22**, 2331216518781746, doi: `10.1177/2331216518781746` pMID: 29956589.

Wilson, R., and Collins, A. (**2019**). "Ten simple rules for the computational modeling of behavioral data," eLife **9**, doi: `10.7554/eLife.49547`.

Bibliography

Yin, T. C., and Chan, J. C. (**1990**). "Interaural time sensitivity in medial superior olive of cat," Journal of Neurophysiology **64**(2), 465–488, doi: `10.1152/jn.1990.64.2.465`.

Yin, T. C., Smith, P. H., and Joris, P. X. (**2019**). *Neural Mechanisms of Binaural Processing in the Auditory Brainstem*, 1503–1575 (John Wiley & Sons, Ltd), doi: `10.1002/cphy.c180036`.

Yost, W. A. (**1974**). "Discriminations of interaural phase differences," J. Acoust. Soc. Am. **55**(6), 1299–1303, doi: `10.1121/1.1914701`.

Yost, W. A. (**1981**). "Lateral position of sinusoids presented with interaural intensive and temporal differences," The Journal of the Acoustical Society of America **70**(2), 397–409, doi: `10.1121/1.386775`.

Zhang, X., Heinz, M. G., Bruce, I. C., and Carney, L. H. (**2001**). "A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression," The Journal of the Acoustical Society of America **109**(2), 648–670, doi: `10.1121/1.1336503`.

Zilany, M. S. A., and Bruce, I. C. (**2006**). "Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery," The Journal of the Acoustical Society of America **120**(3), 1446–1466, doi: `10.1121/1.2225512`.

Zilany, M. S. A., Bruce, I. C., and Carney, L. H. (**2014**). "Updated parameters and expanded simulation options for a model of the auditory periphery," The Journal of the Acoustical Society of America **135**(1), 283–286, doi: `10.1121/1.4837815`.

Zilany, M. S. A., Bruce, I. C., Nelson, P. C., and Carney, L. H. (**2009**). "A phenomenological model of the synapse between the inner hair cell and auditory nerve: Long-term adaptation with power-law dynamics," The Journal of the Acoustical Society of America **126**(5), 2390–2412, doi: `10.1121/1.3238250`.

Zwicker, E. (**1974**). "On a psychoacoustical equivalent of tuning curves," in *Facts and Models in Hearing*, edited by E. Zwicker and E. Terhardt, (Communication and cybernetics), Springer Verlag, pp. 132–141.

Zwislocki, J., and Feldman, R. S. (**1956**). "Just noticeable differences in dichotic phase," J. Acoust. Soc. Am. **28**(5), 860–864, doi: `10.1121/1.1908495`.

List of Publications

Journal paper

Klug, J., Schmors, L., Ashida, G., and Dietz, M. (**2020**). Neural rate difference model can account for lateralization of high-frequency stimuli , The Journal of the Acoustical Society of America **148**(2), 678–691, doi: `10.1121/10.0001602`.

Klug, J., and Dietz, M. (**2022**). Frequency dependence of sensitivity to interaural phase differences in pure tones , The Journal of the Acoustical Society of America **152**(6), 3130–3141, doi: `10.1121/10.0015246`.

Klug, J., Encke, J., and Dietz, M. (**2023**). Characterization of the decline in the auditory nerve phase locking at high frequencies , JASA Express Letters **3**(7), 074403, doi: `10.1121/10.0020267`.

Hu, H., Klug, J., and Dietz, M. (**2022**). Simulation of ITD-dependent single-neuron responses under electrical stimulation and with amplitude-modulated acoustic stimuli , Journal of the Association for Research in Otolaryngology **23**(4), 535–550, doi: `10.1007/s10162-021-00823-1`.

Hu, H., Ausili, S. A., Williges, B., Klug, J., Felsheim, R. C., Vickers, D., and Dietz, M. (**2023**). A model framework for simulating spatial hearing of bilateral cochlear implant users , Acta Acust. **7**, 42, doi: `10.1051/aacus/2023036`.

Klug, J., Heeringa, A., Köppl, C., and Dietz, M. (**2024**). Notched noise improves on-frequency ITD_{ENV} encoding and AN phase locking , in preperation .

Conference abstracts

Klug, J., Schmors, L., Ashida, G., and Dietz, M. (**2-2019**). A physiologically plausible model for human sound lateralization based on envelope interaural time and level differences , 42nd ARO Annual Mitwinter meeting, Baltimore, USA.

Klug, J., Schmors, L., Ashida, G., and Dietz, M. (**09-2019**). Human perception of dichotic high-frequency complex sounds simulated with a two-channel count comparison model in , *23. International Congress on Acoustics*, Aachen, Germany, doi: `10.18154/RWTH-CONV-239637`.

List of Publications

Klug, J., Encke, J., Ashida, G., and Dietz, M. (**1-2020**). Human discrimination of binaural cues in high-frequency sounds with 2-channel model , 43rd ARO Annual Mitwinter meeting, San Jose, USA.

Klug, J., and Dietz, M. (**2-2022**). Human sensitivity to interaural phase difference declines more abruptly than previously thought , 45th ARO Annual Mitwinter meeting, online.

Klug, J., and Dietz, M. (**3-2022**). Human sensitivity to interaural phase difference declines more abruptly than previously thought , 48. DAGA, Stuttgart, Germany.

Klug, J., Heeringa, A., Köppl, C., and Dietz, M. (**2-2023**). Notched noise improves on-frequency ITD$_{ENV}$ encoding , 46th ARO Annual Mitwinter meeting, Orlando, USA.

Heinermann, H. T., Klug, J., Herrmann, S., Ashida, G., Encke, J., and Dietz, M. (**2-2019**). On the frequency limit of interaural time difference sensitivity for pure tones in , *23. International Congress on Acoustics*, Aachen, Germany, doi: `10.18154/RWTH-CONV-239106`.

Heinermann, H., Klug, J., Herrmann, S., Ashida, G., Encke, J., and Dietz, M. (**1-2020**). Frequency limit of itd sensitivity in normal hearing and hearing impaired systems – experimental data and mode , 43rd ARO Annual Mitwinter meeting, San Jose, USA.